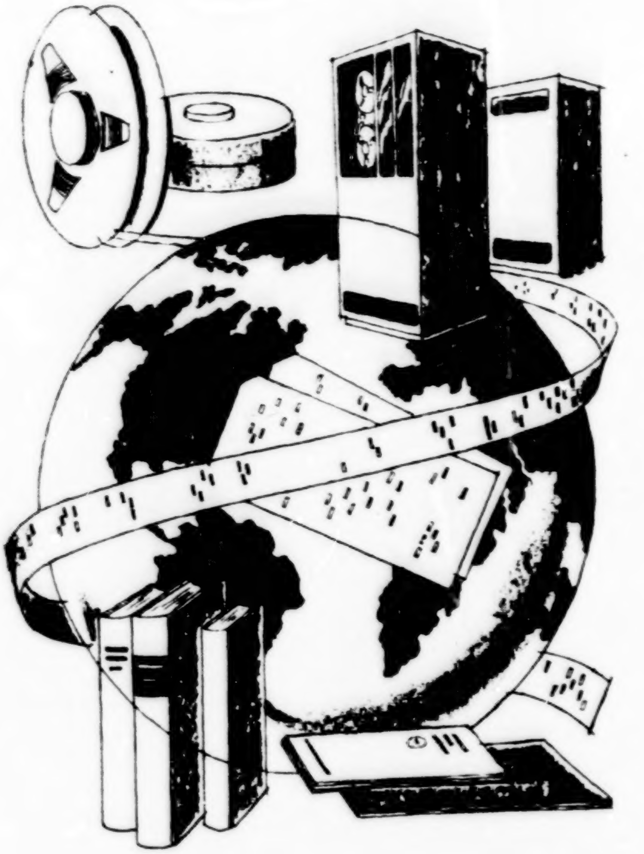
DEC 13 1983

834-e NS1.2:AW1/13/978-980



ORIGINAL

COMPLETED
59.

SUMMARY OF AWARDS

FY 1978, FY 1979, and FY 1980

NATIONAL SCIENCE FOUNDATION
Directorate for Biological, Behavioral, and Social Sciences
Division of Information Science and Technology
Washington, D.C. 20550

CONTENTS

Page

Introduction ii
Definitions iii
FYs 1978, 1979, and 1980 Awards 1
Summaries of Awards 13
Subject Index 51

INTRODUCTION

This report summarizes the awards made by the National Science Foundation in Fiscal Years 1978-1979-1980 on recommendation by the Division of Information Science and Technology.

The report is divided into three major sections. The first section presents a summary table of awards and is arranged alphabetically by principal investigator. The second section, containing descriptions of the work funded, is arranged alphabetically by institutions performing the work. The third section contains the subject index.

Requests for additional information about specific projects should be addressed to principal investigators. General information on the current NSF information science research program and award procedures may be obtained from the Division of Information Science and Technology, National Science Foundation, Washington, DC 20550.

DEFINITIONS

The following definitions are essential to an understanding of this report:

- "Award" refers to financial support given through a grant, contract, interagency agreement, or other arrangement, depending upon the nature of the work to be completed, the terms of performance, and the institution that performs the work.
- "Amendment" refers to an award which adds funds to an existing award with or without increasing the duration of NSF support.
- "Principal Investigator" refers to the scientist or administrator who is responsible for directing an NSF supported project.
- "Institution" refers to any organization which has accepted an award from NSF.

Awards excluded from this report are:

- Purchase Orders
- Personnel Services

The award data presented consist of program grants and contracts made from the Salaries and Expenses appropriations to the Foundation for Fiscal Years 1978-1980. Information contained in the report was obtained from program sources and may differ from official NSF records which may have different inclusions and exclusions.

FY 1978 AWARDS

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Beresford, John	Data Use and Access Labs, Inc.	Conference on a National-Level Program of Access to Machine-Readable Data Files	SOC77-20334	$ 68,300	12/12/77 12	034
Bernard, Russell	West Virginia Univ.	Accuracy of Data on Information Exchange by Electronic Network	IST78-12801	$ 39,403	08/14/78 12	156
Bommer, Michael Chorba, Ronald	Clarkson College of Technology	Analysis of the Operational Parameters for the Management	IST78-12093	$ 76,152	08/30/78 12	026
Caruso, Elaine	U/Pittsburgh	Training Modules for Use of Scientific & Technical Information Services Phase II	IST77-26525	$ 67,875	02/23/79 12	125
Copeland, Richard	Bethune-Cookman College	Measurement and Improvement of On-Line Chemical Searching Performance in Three User Groups at Small Colleges	IST78-11407	$ 89,535	06/13/78 18	007
Coward, H. Roberts	Inst. for Scientific Information	Research Needs and Applications for Indicators Based on the Scientific and Technical Literature	IST78-18730	$ 18,142	09/11/78 6	063
Debons, Anthony	U/Pittsburgh	A Survey of Manpower Requirements for Scientific & Technical Communication	IST77-27115	$101,417	06/26/78 7	126
Dunn, Donald	Stanford Univ.	Economics of Computer Communication Networks	MCS78-02272	$150,000	08/03/78 12	140
Emery, James	EDUCOM	Factors Influencing Network Sharing of Computer Resources in Research	MCS78-07383	$362,269	09/30/78 20	037
Goffman, William	Case Western Reserve Univ.	Theoretical and Experimental Investigation of the Dynamics of Information	IST78-06292	$ 71,588	04/12/78 24	024
Hall, Homer	Rutgers Univ.	Services in Industry for the Analysis and Evaluation of Information	IST78-16628	$ 49,624	08/30/78 24	138
Harris, Virginia	Oberlin College	Research on the Impact of a Computerized Circulation System on the Performance of a Large College Library	IST78-10821	$139,980	09/21/78 24	114
Harris, Zellig	U/Pennsylvania	Correlation of Language Structure with Information	IST78-03863	$163,329	04/10/78 24	120
Hessman, John	Electronic Industries Association	Electronic Information Exchange in Microcomputer Component Standardization Research	IST78-09189	$ 77,113	07/24/78 18	038
Isenhour, Thomas	U/North Carolina	Computer Applications to Analytical Chemistry	CHE78-00632	$ 71,000	05/15/78 12	110

FY 1978 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Kent, Allen	U/Pittsburgh	The Economics of Information Transfer Using Resource Sharing Networks-Network Functions	IST77-17635	$ 84,900	01/10/78 18	130
King, Donald	King Research, Inc.	Development of Statistical Indicators of Scientific and Technical Communication	IST78-04789	$ 9,281	11/15/77 0	068
King, Donald	King Research, Inc.	Data Gathering for Statistical Indicators of Scientific and Technical Communication from Primary Sources	IST77-17915	$147,442	11/22/77 12	069
King, Donald	King Research, Inc.	Systems Analysis of Scientific and Technical Communication in the United States	IST77-28438	$ 11,394	12/08/77 6	070
King, Donald	King Research, Inc.	Evaluation of Information Science Research	IST78-21131	$ 30,000	09/30/78 6	071
Kochen, Manfred	U/Michigan	Links of Interest and Expertise Among Scientists	IST78-16629	$ 49,621	08/14/78 12	091
Lancaster, F. W.	U/Illinois	Impact of a Paperless Society on the Research Library of the Future	IST78-04768	$ 50,540	06/11/78 12	051
Levitan, Karen	Mitre Corporation	Systematic Process Evaluation of Fundamental Information Science Research	IST78-21078	$ 29,847	06/11/78 6	051
Machlup, Fritz	New York Univ.	An In-Depth Analysis of a Comprehensive Data Base on the Communication of Scientific/Scholarly Knowledge	IST78-13383	$160,178	07/10/78 12	104
Martin, Thomas Danowski, James	U/Southern California	Design for Evaluating the Centrality of Basic Information Science Research to Research Users in Organizations	IST78-21130	$ 28,589	08/23/78 6	139
Meadow, Charles	Drexel Univ.	The Design & Assessment of Individualized Instruction in Bibliographic Data Base Access	IST77-26524	$229,600	02/13/78 24	036
Mick, Colin	Applied Communication Research	Development & Testing of a Methodology to Assess Private Investment for Scientific and Technical Information	IST78-10532	$103,350	02/13/78 12	003
Mick, Colin	Applied Communication Research	Assessing the Information Behavior of Scientists & Engineers-A Secondary Data Analysis	IST78-10531	$ 69,309	06/29/78 12	004
McGill, Michael	Syracuse Univ.	An Evaluation of Factors Affecting Document Ranking by Information Retrieval/Systems	IST78-10454	$ 67,857	05/09/78 12	148

FY 1978 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Niehoff, Robert	Battelle Memorial Inst.	Extension, Refinement and Evaluation of On-Line Automated Subject Switching	IST77-04498	$106,186	05/26/78 12	006
Nisenoff, Norman	Forecasting International	A Study of the Potential Influence of Social, Economic, Regulatory & Technological Factors on Scientific & Technical Communication Through 2000 A.D.	IST78-12102	$178,775	09/13/78 24	042
O'Connor, John	Lehigh Univ.	Uses of Citing Statements to Improve Retrieval	IST78-01054	$116,037	07/05/78 24	073
Olson, Edwin Lake, Dale	NTL Inst. for Applied Behavioral Science	Development of a Conceptual Framework to Guide Research in Information Science	IST78-19923	$ 49,611	09/07/78 12	112
Paisley, William	Stanford Univ.	Developing a Model and Methodology for Evaluating Practitioner Information Systems: The Case of Pacfornet	IST77-28242	$ 65,000	09/21/78 12	144
Riley, James	Library of Congress	Evaluation of the Use of Slow-Scan Television and Telefacsimile in a Library Network	IST78-12101	$145,905	08/23/78 24	076
Rummer, Dale	U/Kansas	Improving the Use of Computer-Based Information Resources Within the Engineering Classroom	IST77-23193	$ 61,421	01/31/78 15	067
Sager, Naomi	New York Univ.	Research into Methods for Automatic Classification & Fact Retrieval in Science Subfields	IST77-24530	$211,000	12/29/77 24	107
Sager, Naomi	New York Univ.	Data Base Management Systems for Natural Language Data Bases	IST78-13939	$188,943	09/07/78 24	108
Salton, Gerard	Cornell Univ.	Research on Combining Document and Data Retrieval Technologies	IST77-04843	$133,555	11/18/78 24	030
Saltzer, Benjamin	Naval Ocean Systems Center	Study of Impacts of a Formal Computer-Based Information System on Information Networks	IST77-23192	$ 71,891	05/23/78 18	099
Schipma, Peter	IIT Research Institute	Video Full Text Processing	IST78-11381	$222,570	08/23/78 18	049
Sheridan, Thomas	Massachusetts Institute of Technology	Electronic Information Exchange in Mental Workload Research	IST78-06749	$148,454	07/19/78 18	085
Shuchman, Hedvah	The Futures Group	The Effects of STI Changes on the Engineering Profession	IST77-20839	$129,927	06/13/78 15	044

FY 1978 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Small, Henry Pinski, Gabriel	Inst. for Scientific Information	The Structure & Development of the Information Science Literature: A Citation Study	IST78-16677	$ 32,689	08/23/78 12	062
Stevens, C. H.	Participation Systems	Electronic Information Exchange in Legislative Technical Research	IST77-27940	$150,575	07/24/78 18	119
Turoff, Murray	New Jersey Institute of Technology	Electronic Information Exchange Test Facility	IST77-21008	$275,197	08/03/78 18	101
Waltz, David	U/Illinois	Theoretical Issues in Natural Language Processing-2: An Interdisciplinary Conference	IST78-18228	$ 14,242	06/26/78 6	052
White, Herbert	Indiana Univ.	The Economics of Library & Subscriptions: A Sample Survey	IST78-09182	$ 61,025	06/13/78 12	059
Wilcox, Alice	Minnesota Higher Educ. Coordinating Board	A Conference on Serial Collections in a Network Environment	IST78-18103	$ 32,469	09/29/78 12	093
Williams, Martha	U/Illinois	An Integrated Man/Machine Interface to Facilitate Network Resource Utilization	IST77-26886	$174,432	02/14/78 18	055
Zinn, Karl	U/Michigan	Computer-Based System for Sharing Technical Information	IST78-03962	$ 20,000	05/05/78 12	089

FY 1979 AWARDS

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Agrawal, Dharma	Wayne State Univ.	Subgraph Identification Using Associative Techniques with Applications to Information Science and Chemical Structure Investigation	IST79-18462	$ 61,470	07/30/79 24	154
Allen, Beth Elaine	U/Pennsylvania	Alternative Representations of Information in Microeconomic Systems	IST79-18464	$ 77,987	07/30/79 24	121
Braunstein, Yale	Brandeis Univ.	Economics of Market Structure in Scientific and Technical Information Activities	IST79-15235	$156,523	09/06/79 24	009
Chapanis, Alphonse Weeks, Gerald	Johns Hopkins Univ.	Human Requirements in Information Retrieval	IST78-20923	$195,841	02/05/79 24	066
Cohen, Karen	Massachusetts Institute of Technology	Experimental Evaluation of Information Delivered in Various Modes and Combination of Modes	IST78-20362	$131,746	09/26/79 24	082
Cooper, Franklin	Haskins Laboratories	Verification of Syllable Hypotheses in Automatic Phonetic Analysis of Spoken English	MCS79-16177	$109,475	08/30/79 12	047
Cooper, William	U/California, Berkeley	Bayesian Retrieval System Design	IST79-17566	$ 97,125	08/31/79 24	013
Cozzens, Susan	Inst. for Scientific Information	User Requirements for Scien- Reviews	IST78-21947	$ 31,421	11/07/78 12	064
Debons, Anthony	U/Pittsburgh	A Survey of Manpower Requirements for Scientific & Technical Communication	IST79-09524	$ 20,440	04/02/79 12	127
Debons, Anthony	U/Pittsburgh	A Survey of Manpower Requirements for Scientific & Technical Communication	IST79-18769	$126,874	06/20/79 12	128
Debons, Anthony	U/Pittsburgh	Support of U.S. Participation in NATO Advanced Study Institute in Information Science	IST78-16989	$ 15,189	08/23/79 12	129
Dunn, Donald	Stanford Univ.	Economics of Computer Communication Networks	MCS78-02272	$157,172	07/01/79 12	141
Evens, Martha	Illinois Inst. of Technology	Structuring the Lexicon and the Thesaurus with Lexical-Semantic Relations	IST79-18467	$ 43,575	08/06/79 24	058
Flowers, John	U/Nebraska	Human Attention and Performance Limitations in Processing Visual Information	IST79-11162	$ 34,672	07/30/79 12	100
Friedman, Joyce	U/Michigan	Computer Studies in Formal Linguistics	BNS79-05879	$ 87,508	06/28/79 24	090

FY 1979 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Gell, Marilyn	National Commission on Libraries and Information Science	Electronic Information Exchange in Planning for White House Conference on Library and Information Services	IST79-21035	$ 40,335	08/27/79 7	098
Gerstenfeld, A. Franke, Richard	Worcester Polytechnic Institute	The Critical Review Process: Impact Upon Scientific and Technological Information	IST79-16483	$147,255	08/24/79 24	159
Hammond, Kenneth Cook, Richard Taylor, Robert	U/Colorado	Designing Information Systems for More Effective Use in Decision Making	IST78-19867	$105,531	05/02/79 12	027
Hayes, Robert Borko, Harold	U/California, Los Angeles	Mathematical Models of Information System Use	IST79-18497	$196,640	08/14/79 24	017
Hickey, Thomas	OCLC, Inc.	The Probabilistic Matching and Control of Author Names in Automated Library Systems	IST79-18263	$ 42,321	08/21/79 12	115
Joshi, Aravind	U/Pennsylvania	Computer Science and Computer Engineering Research Equipment	MCS79-07558	$175,100	07/02/79 12	122
Kent, Allen	U/Pittsburgh	Economics of Information Transfer Using Resource Sharing Networks—Network Modeling Simulation	IST79-11886	$233,496	08/13/79 30	131
Kidd, Jerry	U/Maryland	Research on Behavioral Variables Related to Information Systems	IST79-04896	$163,011	08/22/79 18	079
Langley, Patrick	Carnegie-Mellon Univ.	An Information Processing Theory of Procedural Learning	IST79-18266	$ 53,715	07/30/79 24	022
Lehnert, Wendy Schank, Roger Abelson, Robert P.	Yale Univ.	Text Processing Effects and Memory Recall	IST79-18463	$291,988	08/03/79 24	160
Loy, James	National Bureau of Standards	Support of the Numerical Data Advisory Board	IST79-20632	$ 12,300	08/30/79 12	097
Maier, David	SUNY, Stony Brook	Relational Theory of Databases	IST79-18264	$ 89,655	08/08/79 24	146
McCarroll, Jane	Innovative Research, Inc.	Electronic Information Exchange in Research on Devices for the Disabled	IST79-18251	$ 20,275	06/06/79 12	061
Michael, Elizabeth	Tymshare, Inc.	Innovations in Engineering Publications	IST79-06604	$ 3,695	12/15/78 24	152
Niehoff, Robert	Battelle Memorial Inst.	Advanced Concepts and Studies of a Vocabulary Switching System	IST79-11190	$148,900	09/21/79 12	005
Parsons, Charles Levi, Isaac	Columbia Univ.	Normal Sentence Forms and Information Structure	IST79-19203	$354,377	09/06/79 24	029

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Pearl, Judea	U/California, Los Angeles	Investigating Computational Gains Using Partial Information	MCS78-18924	$ 68,751	03/09/79 12	018
Perlmutter, Jane	U/Texas, Austin	Representation, Retrieval and Reorganization of Conceptual Information	IST79-18474	$ 63,952	07/25/79 24	150
Pollock, Joseph	American Chemical Society	Automatic Spelling Error Detection and Correction for Large Data Bases	IST78-20175	$153,160	05/31/79 30	002
Reintjes, J. Francis	Massachusetts Inst. of Technology	A Research Program on Electronic Document-Delivery Systems for Interlibrary Loans	IST78-23251	$ 83,725	02/14/79 12	084
Roderer, Nancy Wiederkehr, Robert	King Research, Inc.	Evaluation Measures and Methods for On-Line Bibliographic Systems	IST79-18479	$143,864	08/14/79 15	072
Salton, Gerard	Cornell Univ.	Mathematical Modelling in Automatic Information Retrieval	IST79-05301	$145,040	05/03/79 24	031
Salton, Gerard	Cornell Univ.	Fast Computational Processes in Information Retrieval	IST79-09176	$ 29,605	07/27/79 12	032
Slamecka, Vladimir	Georgia Inst. of Technology	Syntactic Structure of Information and Information Processing	IST78-27002	$137,790	08/01/79 24	045
Sullo, Pasquale Wallace, William	Rensselaer Polytechnic Institute	A Reliability Theoretic Construct for Document Usage and Information Flow	IST78-16414	$ 72,800	06/22/79 12	135
Tannebaum, Aaron	Brooklyn College	Efficiency of Sequential Search Algorithms	IST79-17568	$ 38,376	07/27/79 12	010
Tou, Julius	U/Florida	Development of Theory and Techniques for Automated Knowledge Utilization, Transfer & Extension	IST79-08548	$166,740	08/03/79 24	039
Tsao, Yao-chung	Rensselaer Polytechnic Institute	Information Processing of Logographic and Alphabetic Writing Symbols	IST79-18582	$ 27,850	08/06/79 12	136
Webber, Bonnie	U/Pennsylvania	Research in Natural Language Man-Machine Interaction	IST79-18476	$ 83,598	08/01/79 24	123
Werdel, Judith	National Academy of Sciences	International Scientific and Technical Information Programs	INT79-02931	$ 96,500	03/06/79 12	095
Wood, James Frase, Robert	Council of National Library Association	ANSI Committee Z39	IST78-25323	$ 20,000	05/10/79 12	033
Yovits, Marshall	Ohio State Univ.	Extension & Applications of a Theory of Information Flow and Analysis: The Use of Information in a Decision-Making Environment	IST79-08327	$105,670	07/16/79 24	117

FY 1979 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Zadeh, Lotfi	U/California, Berkeley	A Theory of Approximate Reasoning	MCS79-06543	$ 29,000	08/09/79 12	014
Zunde, Pranas	Georgia Institute of Technology	Study of Potential New Applications of Information Theory to Problems of Information Science	IST79-17567	$ 82,449	09/10/79 12	046

FY 1980 AWARDS

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Allen, James F.	U/Rochester	Recognizing Intent in Natural Language	IST80-12418	$ 80,000	08/15/80 24	137
Anderson, James	Brown Univ.	Cognitive Applications of Matrix Memory Models	BNS79-23000	$ 44,602	03/31/80 12	011
Becker, David S.	IIT Research Inst.	An Augmented English Dictionary For Information Retrieval	IST79-20267	$ 83,285	06/26/80 12	050
Bernard, H. Russell	U/Florida	Degradation of Accuracy in Social Network Data	IST79-24965	$ 19,403	12/18/79 12	040
Brown, Donna J.	U/Illinois, Urbana	Storage and Access Costs in Data Structures for Information Retrieval Problems	IST80-12240	$ 62,663	08/14/80 24	053
Chase, Karen H.	Pennsylvania State Univ.	Completions of Relations and Embedded Join Dependencies in Relational Databases: A Study Using Boolean Matrices	IST80-12239	$ 36,630	06/26/80 12	124
Chu, Wesley W.	U/California, Los Angeles	File Allocation and Query Optimization for Distributed Data Bases	IST80-02671	$112,220	07/17/80 24	019
Croft, W. Bruce	U/Massachusetts, Amherst	A Generalized Model of Document Retrieval	IST80-11605 24	$ 61,740	06/27/80	081
Dunn, Donald	Stanford Univ.	Economics of Computer Communication Networks	MCS80-12875	$166,458	08/01/80 12	142
Eastman, Caroline	Florida State Univ.	Multiple Attribute Retrieval Techniques for Tree-Structured Information	IST80-11608	$ 24,063	08/07/80 12	041
Ellis, Carla S.	U/Oregon	Concurrency in the Access and Manipulation of Information Structures	IST80-12441	$ 62,500	06/18/80 24	118
Falmagne, Jean-Claude	New York Univ.	Probabilistic Measurement Theoretic Models of Information Processing	IST79-24526	$156,909	04/02/80 31	103
Fu, King-Sun	Purdue Univ.	Research on Error-Correcting Techniques in Automated Information Processing	IST79-18884	$151,794	11/05/80 36	132
Grossberg, Stephen	Boston Univ.	Decision-Making and Information Processing in Real-Time Network Models	IST80-00257	$157,990	04/02/80 36	008
Harris, Zellig	Columbia Univ.	Information Correlates of Basic Language Structures	IST80-17043	$ 74,989	09/23/80 24	028
Haseman, William D.	Carnegie-Mellon Univ.	Distributed Data Base Design	IST79-18472	$ 39,211	12/07/79 6	023
Haseman, William D.	U/Wisconsin, Milwaukee	Distributed Data Base Design	IST80-19259	$243,651	08/04/80 18	158

FY 1980 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Hillman, Donald	Lehigh Univ.	A Theoretical Basis for Knowledge Management	IST80-12234	$123,551	06/19/80 24	074
Hiltz, S. Roxanne	Upsala College	Workshop to Compare and Synthesize Evaluations of Computer-Based Information Exchange	IST80-18077	$ 43,711	06/10/80 6	153
Huber, George	U/Wisconsin, Madison	A Study of the Distribution, Processing, and Use of Scientific and Technical Information in Research Organizations	IST79-13382	$ 91,214	01/23/80 12	157
Isenhour, Thomas	U/North Carolina	Computer Applications to Analytical Chemistry	CHE80-40035	$ 66,500	06/11/80 12	111
Kaske, Neal K.	OCLC, Inc.	Development of Generalized Algorithm to Determine Terminal Requirements for On-Line Catalogs in Libraries	IST79-20312	$ 97,902	05/22/80 12	116
Knuth, Donald E.	Stanford Univ.	Theoretical Basis for the Development of Document Preparation Systems	IST79-21977	$338,707	01/23/80 24	143
Kochen, Manfred	U/Michigan	Dynamics in the Organization of Scientific Specialities	IST80-07433	$134,782	08/07/80 24	092
Koll, Matthew B.	National Bureau of Standards	Refinement and Use of Concept Similarities to Information Retrieval	IST80-11607	$ 35,498	09/01/80 12	096
Leimkuhler, Ferdinand F.	Purdue Univ.	Analysis and Application of Information Productivity Models	IST79-11893	$108,099	01/16/80 24	133
Leimkuhler, Ferdinand F.	Purdue Univ.	Analysis and Application of Information Productivity Models	IST80-12834	$ 1,819	06/06/80 0	134
Machlup, Fritz	New York Univ.	Information Science: An Analysis of Methodological Issues and Interdisciplinary Relationships	IST80-08007	$ 39,973	05/27/80 12	105
Marcus, Richard S. Reintjes, J. Francis	Massachusetts Institute of Technology	Investigations of Computer-Aided Document Search Strategies	IST80-06516	$155,360	06/11/80 24	086
McDonough, Carol	U/Lowell	A Simultaneous Equation Model of the Demand for Academic and Professional Journal	IST79-19478	$ 25,342	05/09/80 12	078
McGill, Michael	Syracuse Univ.	A Study of the Impact of Representations in Information Retrieval Systems	IST79-21468	$155,613	02/05/80 24	149
Minker, Jack Jacobs, Barry	U/Maryland	Applications of Logic and Theorem Proving to Databases	MCS79-19418	$207,835	04/18/80 36	080

FY 1980 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Moss, Mitchell L.	New York Univ.	Publication of Proceedings of a Conference on Telecommunications and Productivity	IST80-05145	$ 9,976	04/08/80 12	106
Mount, Kenneth	Northwestern Univ.	Message Spaces and Computing Time in Performance Allocation Processes	IST80-07457	$ 78,409	08/07/80 24	113
Moyne, John	Queens College of CUNY	Language and Human Information Processing	IST80-23585	$ 91,140	05/30/80 15	025
Narens, Louis E.	U/California, Irvine	Problems in the Theory of Measurement that Pertain to Information Concepts	IST79-24019	$111,264	03/11/80 36	016
Neff, Karen J.	Hofstra Univ.	Research on Automated Natural Language Generation	IST80-06945	$110,633	08/07/80 24	048
Noma, Elliot	Institute for Scientific Information	Citation-Based Measures of Interactivity and Similarity	IST79-23661	$ 38,188	03/06/80 12	065
O'Connor, John	Lehigh Univ.	Citing Statements: Recognition by Computer and Use to Improve Full-Text Retrieval	IST80-04353	$115,447	07/22/80 24	075
Parker, D. Stott	U/California, Los Angeles	Data Semantics and Logical Database Design	IST80-12419	$ 57,489	07/30/80 24	021
Pearl, Judea Dalkey, Norman	U/California, Los Angeles	Studies in the Organization of Information Structures	IST79-22062	$ 39,174	03/21/80 18	020
Peters, Alec	Franklin Institute Research Laboratories	Support for Coordination of Scientific and Technical Information Activities	IST79-24639	$ 50,000	12/04/79 7	043
Pierrehumbert, Janet	Massachusetts Institute of Technology	The Use of Intonation in Automatic Speech Understanding	IST80-12248	$ 43,701	07/03/80 24	083
Puri, Madan L.	Indiana Univ.	Problems in Decision-Making and Information Retrieval Studied with Techniques of Fuzzy Sets and Systems	IST79-18468	$112,690	12/05/79 36	060
Reintjes, J. Francis	Massachusetts Institute of Technology	Investigations of Electronic Interlibrary Resource-Sharing Networks	IST80-1664	$ 95,000	06/11/80 12	087
Rich, Elaine	U/Texas, Austin	Individual User Modeling in Computer Systems	IST80-11614	$ 56,315	08/04/80 24	151
Richards, Whitman	Massachusetts Institute of Technology	Natural Computation and Control	MCS79-23110	$234,196	04/11/80 12	088
Rouse, William B.	U/Illinois, Urbana/ Champaign	Modeling of Human Behavior in Seeking & Generating Information	IST79-16471	$144,120	12/21/79 24	054
Sager, Naomi	New York Univ.	Computer-Based Investigation into the Structure of Information	IST79-20788	$274,645	01/25/80 24	109

FY 1980 AWARDS—CONTINUED

DIVISION OF INFORMATION SCIENCE AND TECHNOLOGY

Principal Investigator(s)	Institution	Title	Award No.	Amount	Award Date/ Duration (Month)	Summary No.
Schank, Roger	Yale Univ.	Theoretical Foundation of Natural Language Processing	IST80-17790	$183,785	09/17/80 24	161
Srihari, Sargur N.	SUNY, Buffalo	Contextual Algorithms for Text Recognition	IST80-10830	$ 55,800	07/10/80 24	145
Swamy, Sowmitri	U/Illinois, Urbana	Access Time Versus Storage Space in Information Retrieval Systems	IST80-12242	$ 59,829	06/26/80 24	057
Tenenhaus, Michael	Wayne State Univ.	Lexical Ambiguity Resolution in Discourse	IST80-12439	$38,977	08/25/80 24	155
Turoff, Murray	New Jersey Institute of Technology	Electronic Information Exchange Test Facility	IST80-09260	$ 2,000	03/28/80 0	102
Ullman, Jeffrey	Dept. of Air Force	Implementation of a Simple User Model by a Relational Database	IST80-21358	$ 49,487	08/12/80 24	001
Wainer, Howard	Bureau of Social Science Research	Social Science Applications of Dynamic Multivariate Displays	SES80-08481	$ 40,728	05/21/80 12	012
Warren, David	SUNY, Stony Brook	Montague Grammar & Database	IST80-10834	$ 79,902	08/06/80 24	147
Weischedel, Ralph	U/Delaware	Natural Language Processing Systems and Ill-Formed Input	IST80-09673	$136,681	08/04/80 24	035
Whitehead, Bruce	U/Louisville	Representation and Processing of Spatially Structured Information in Cognition	IST80-11617	$ 77,840	09/05/80 24	077
Williams, Martha	U/Illinois, Urbana/ Champaign	Automatic Duplicate Detection of Journal Articles Appearing in Multiple Data Bases—A Feasibility Study	IST79-21018	$269,303	02/13/80 24	056
Zadeh, Lotfi A.	U/California, Berkeley	A Theory of Approximate Reasoning	MCS79-06543	$ 29,500	04/18/80 12	015

SUMMARIES OF AWARDS
FISCAL YEARS 1978, 1979, 1980

#001

Implementation of a Simple User Model by a Relational Database. *Department of the Air Force,* Bolling Air Force Base, Washington, DC 20332. Jeffrey Ullman. $49,487 for 24 months beginning August 12, 1980.

It is proposed to build an information retrieval system based on relational database theory, which would be both efficient and easy to use.

It is a great convenience for the user if a database can be viewed as a single (*universal*) table, and queries are posed about this imaginary table. However, it is generally too expensive of storage space to allow the redundance that would necessarily be present in such a table. Thus, relational databases usually consist of several tables, and queries are posed in terms of combinations of these tables. This underlying database structure should be made invisible to the user, but several problems are encountered. One is the design problem: what tables should be used to store the database? A second is query interpretation: what access paths (connections between data in different tables) are implied in queries about the imaginary universal table? A third is optimization: how to optimize the query on the underlying tables? A fourth problem to consider is how updates are to be performed. The user may attempt to insert or delete associations from the universal table that are not reflected naturally in the underlying tables.

The project, in brief, will produce a system capable of (a) interpreting queries on the universal table as if they were queries on the underlying database and (b) assisting the database designer in selecting the underlying tables and in specifying the rules whereby query interpretation can be conducted correctly. Appropriate theory will be developed where needed.

#002

Automatic Spelling Error Detection and Correction for Large Data Bases. *American Chemical Society,* P. O. Box 3012, Columbus, OH 43210. Joseph J. Pollock. $153,160 for 30 months beginning May 31, 1979.

Chemical Abstracts Service (CAS) will undertake a 30-month research program in automatic detection and correction of spelling errors in free text. The proposed research has value for broader questions than the correction of spelling errors in the CAS file; it is expected to make available to other researchers a data base and proven methods which can be generalized to a variety of significant information science questions concerned with the identification of semantic content in terms of syntactic form, including applications to devices which produce text outputs in response to speech inputs.

Data on spelling errors will be gathered. Word classification procedures to identify and separate vocabularies (e.g., scientific English, taxonomic terms, and systematic chemical names) will be investigated, as will word structure analysis techniques. The applicability of n-gram analysis to spelling error detection will be determined.

A dictionary of repeated misspellings in the experimental corpora will be compiled and the proportion of the total misspellings that can be corrected using this dictionary will be determined. Correction heuristics using co-occurrence data at several levels, and the usefulness of n-gram analysis in conjunction with word classification will be studied.

#003

Development and Testing of a Methodology to Assess Private Investment for Scientific and Technical Information. *Applied Communication Research, Inc.,* P. O. Box 5849, Stanford, CA 94305. Colin K. Mick. $103,350 for 12 months beginning February 13, 1978.

The objective of this project is to develop a conceptual framework to define STI expenditures in industry and to develop and test a methodology for collecting these data. Because of the conceptual and definitional problems, as well as the difficulties of actually providing such data, close working relationships will be established with representatives of all companies to be surveyed. A task force of industry representatives will also be established to provide guidance and input to the researchers. The survey sample will be based upon over 500 companies included in the annual *Business Week* analysis of corporate investment in research and development.

#004

Assessing the Information Behavior of Scientists and Engineers—A Secondary Data Analysis. *Applied Communication Research, Inc.,* P. O. Box 5849, Stanford, CA 94305. Colin K. Mick. $69,309 for 12 months beginning June 29, 1978.

The purpose of this study is to explore the information behavior of scientists and engineers utilizing recently collected data. Since no new data collection is proposed for this project, the emphasis will be on identifying and testing existing information behavior hypotheses and on identifying and exploring relationships between information behaviors and between those behaviors and contributing factors such as individual attitudes, environmental constraints and situational (task) constraints.

Products of this study will include a series of models for predicting information behaviors, an annotated listing of information behavior hypotheses, and an annotated bibliography of the information behavior research literature.

#005

Advanced Concepts and Studies of a Vocabulary Switching System. *Battelle Memorial Institute,* 505 King Avenue, Columbus, OH 43201. Robert T. Niehoff. $148,900 for 12 months beginning September 21, 1979.

There are many bibliographic data bases available for on-line searching. Using these data bases to locate information for an interdisciplinary project is very difficult because each data base uses the specific vocabulary to its subject discipline which may not be familiar to the researcher. A vocabulary switching system (VSS) has been developed which will accept a user's search terms and convert them to the indexing terms used in other fields. This permits the retrieval of documents on a given subject from several data bases with a single query. The VSS, which was originally developed for energy terms, has been extended and now can be used to correlate terms from several large controlled vocabularies. This project will investigate subject switching between controlled and uncontrolled vocabularies. The effects of variables such as the similarity or dissimiliarity of data bases, size and number of data bases, and the similarity or dissimilarity of the subject fields will be evaluated to determine the circumstances under which subject switching provides the greatest benefits. The role of subject switching in an intermediary-free search environment will also be investigated.

#006

Extension, Refinement and Evaluation of On-Line Automated Subject Switching. *Battelle Memorial Institute,* 505 King Avenue, Columbus, OH 43201. Robert T. Niehoff. $106,186 for 12 months beginning May 26, 1978.

On-line searching of multiple data bases for bibliographic records is extremely time consuming and very costly. This is due to the diversity of subject indexing procedures and vocabularies used. These create a barrier to greater utilization of existing information resources. The main purpose of the Battelle project is to develop, test and evaluate a technique which will permit the user to search on-line computerized data bases outside his subject field. The proposed computer procedure will accept the user's own, familiar subject terms, correlate and convert them to the indexing terms used in other subject fields. This will permit the retrieval of documents on a given subject from all available data bases with a single query.

The project's investigator plans to expand his earlier investigations in indexing energy information. Previously, from selected energy related terms used by a group of computerized data bases, he developed a preliminary, on-line subject switching technique for energy. To the original energy vocabulary, he will add new terms, refine the original software, and attempt to determine whether the success obtained in energy can also be achieved in a much larger group of fields.

The result will be a better understanding of how to transfer information from computerized data bases to the end user.

#007

Measurement and Improvement of On-Line Chemical Searching Performance in Three User Groups at Small Colleges. *Bethune-Cookman College,* Daytona Beach, FL 32015. Richard F. Copeland. $89,535 for 18 months beginning June 13, 1978.

The high fixed minimum cost of maintaining a traditional library of bound scientific and technical journals is more than many small undergraduate colleges are capable of supporting. Significant savings have been obtained by substituting on-line searching for abstract journal subscriptions. This suggests that the 1100 to 1300 such institutions may well be using on-line searches as their primary entry into the scientific and technical literature within the next few years. An increasing portion of this on-line bibliographic searching will be done by inexperienced, infrequent users—college faculty and students in the case of undergraduate colleges. The relative proficiency of these users has not been established. Phase I of the project will study the searching techniques and quantify the proficiencies of three distinct groups: information scientists who are full-time on-line searchers, college chemistry faculty who are part-time searchers, and novice searchers. Based on the results of Phase I, Phase II will investigate whether an on-line prompting system can significantly improve the absolute searching efficiency of any or all of the three groups of searchers. Recall, precision, and search time and cost are the variables to be measured. Randomized, complete block designs will be used for standard analyses of variance.

#008

Decision-Making and Information Processing in Real-Time Network Models. *Boston University,* Department of Mathematics and Biomedical Engineering, Boston, MA 02215. Stephen Grossberg. $157,990 for 36 months beginning April 2, 1980.

The proposed research is an investigation of the properties of rapid parallel decision-making and stable self-organization of environmentally sensitive codes that occur in network models which have previously been derived from psychological postulates.

A key project is the study of absolutely stable pattern transformers, including their short term storage, choice behavior, and self-tuning capabilities. Additional work will deal with normative drifts and pattern completion, the functional units of cognitive coding, and oscillatory behavior in parallel processors.

The work will draw upon the investigator's continuing work on a general method for classifying the dynamics of competitive systems. This method, which is based on differential equations, derives from his previous work on the neural basis of cognitive mechanisms. It provides a tool for designing, analyzing, and classifying competitive systems that have desirable biological behavior. The results have potential for application to the representation of knowledge in both human and electronic information processing systems.

#009

Economics of Market Structure in Scientific and Technical Information Activities. *Brandeis University.* Waltham, MA 02154. Yale M. Braunstein. $156,523 for 24 months beginning September 6, 1979.

This research is designed to examine both the market structure of the Scientific and Technical Information (STI) industries and related economic phenomena. The project combines economic theory and empirical testing with the objective of contributing to our understanding of such issues as pricing, market entry, vertical integration and size of firm. Four specific areas of research will be addressed. These include the development of cost functions for firms engaged in different phases of the dissemination of STI; measurement and analysis of technological change over time; examination of the motivation of the firm in the STI industries; and a comprehensive analysis of the structure of the STI industries.

#010

Efficiency of Sequential Search Algorithms. *Brooklyn College,* Department of Computer & Information Science, Brooklyn, NY 11210. Aaron Tennebaum. $38,376 for 12 months beginning July 27, 1979.

This research will examine the efficiency of a class of sequential search algorithms in which an item which is accessed is moved to a forward position on the list and will need fewer searches. The effects of various parameters will be examined. These parameters include the number of elements on the list and number of searches. Optimal points at which to switch from one algorithm to another will be identified.

#011

Cognitive Applications of Matrix Memory Models. *Brown University,* Department of Psychology and Center for Neural Sciences, Providence, RI 02912. John A. Anderson. $44,602 for 12 months beginning March 31, 1980.

A model for memory based on ideas from neuroscience has been developed and published previously. It is assumed that (a) nervous system activity is most usefully represented as a set of simultaneous individual neuron activities in a group of neurons (a state vector) and (b) changes in synaptic strength of a particular (correlational) kind store memory forming a matrix of synaptic connectivities which can couple associated state vectors. Recent extensions of this model involved positive feedback of a set of neurons on itself and introduction of a simple limiting non-linearity. This work will develop this model further. Such models seem to be very useful for describing human perceptual processes quite close to the sensory input. Non-linear extensions, to be investigated, seem to be able to make significant contact with questions of more cognitive relevance such as categorization, logical inference, and semantic networks. Such models may have considerable practical use since they are able to generalize, associate and categorize, yet are immune to certain pathologies of large systems: that is, since there are a fixed number of modifiable synapses operating in parallel, there is little possibility of the kinds of exponential retrieval and storage disasters that haunt data bases of a very large size.

#012

Social Science Applications of Dynamic Multivariate Displays. *Bureau of Social Science Research,* 1990 M Street, NW, Washington, DC 20036. Howard Wainer. $40,728 for 12 months beginning May 21, 1980.

A newly developed method for the interactive display of multivariate data (PRIM-9) will be used to examine diverse sets of social data chosen for their generality of form. The results will be compared to results obtained using static graphical display methods. The aim is to assess the value of the new method for detecting structure and relationships in data of great complexity. Transportable data sets representative of broad classes of social data will be prepared for manipulation at the one existing PRIM-9 facility in the U.S. These data will be drawn from bodies of data widely used in economic, demographic, and social indicators research, and they will include several levels of geographic aggregation, individual-level data taken from field experiments and social surveys, and simulated data of known (constructed) multivariate structure. The PRIM-9 system's utility in revealing complex interrelationships, generating hypotheses about causal structures, and testing explicit theoretical propositions will be examined. The results will be recorded on video tape and written into a description and evaluation of PRIM-9 and PRIM-9-like systems for social science data analysis. Formal and informal means of reporting will be employed to inform the social science research community of the effectiveness of this new technology and to provide feedback to the originators and developers of PRIM-9. Such reports will include suggestions for modifications and improvements in the system which would increase the analytic power available to the research community.

#013

Bayesian Retrieval System Design. *University of California*, School of Library & Information Studies, 118 South Hall, Berkeley, CA 94720. William S. Cooper. $97,125 for 24 months beginning August 31, 1979.

The aim of this research is to develop mathematical and statistical techniques for the design of more effective information retrieval systems. The emphasis is on improving the underlying indexing and search logic used in document and reference retrieval systems rather than on developing more sophisticated hardware.

Retrieval will be treated as a problem in probability estimation. A retrieval system should be able to determine which documents in a collection have the highest probability of usefulness, and to retrieve the documents in that order. Such probabilities can be best estimated on the basis of a combination of several types of data, and the methods of Bayesian statistics will be used to find practical ways of combining such evidence to perform the estimation within the computer. Estimation formulas will be developed for various types of search requests and for a variety of standard and novel statistical clue-types. The mathematical basis of a possible "browsing" capability for advanced systems will also be investigated.

#014

A Theory of Approximate Reasoning. *University of California*, Electronics Research Laboratory, Berkeley, CA 94720. Lotfi A. Zadeh. $29,000 for 12 months beginning August 9, 1979.

This research is aimed at the development of a method of inference from imprecise or unreliable data which are expressed in the form of propositions in a natural or synthetic language. The main application areas for this method lie in the design of expert question-answering systems, decision analysis, and control of robotic systems operating in an imprecisely defined environment.

The approach of this research is based on the representation of the meaning of an imprecise proposition by a procedure which acts on a set of relations in a database and returns a possibility distribution for the variables which are associated with the proposition. The inference rules are applied to the translation of the premises and yield a possibility distribution which upon retranslation becomes an imprecise conclusion inferred from the premises. In general, the premises are allowed to be qualified with imprecisely defined truth values, probability values or possibility values.

#015

A Theory of Approximate Reasoning. *University of California*, Electronics Research Laboratory, Berkeley, CA 94720. Lotfi A. Zadeh. $29,500 for 12 months beginning April 18, 1980.

This supplements work described in the above project.

#016

Problems in the Theory of Measurement that Pertain to Information Concepts. *University of California*, School of Social Sciences, Irvine, CA 92717. Louis E. Narens. $111,264 for 36 months beginning March 11, 1980.

The term "measurement," as used here, refers to the study of scientifically interesting collections of objects or events whose elements exhibit some attribute that can be ordered quantitatively. That is, given any two elements of the collection, it is possible to determine by some well prescribed procedure which of the two has more of the attribute or if they exhibit it to the same degree. From a scientific point of view, the most difficult problem usually is how to device systematic ways of discovering the ordering of the attribute in question.

The purpose of this grant is to work on some mathematical problems that appear to be useful in the measurement and the understanding of aspects of cognition that are relevant to principles of information science. These can be grouped into three major headings having to do with increasingly complex levels of structure: (1) Structures with one operation: concatenation and conjoint ones; (2) Structures with two operations: probability and random variables; (3) Structures with many operations: dimensional analysis and meaningfulness. Such studies of the fundamental structure of information are concerned with issues common to the empirical aspects of all sciences.

The research will include work on some of the important problems in the axiomatic foundations of measurement theory, particularly those that pertain to the concepts of probability, utility, risk, and information.

#017

Mathematical Models of Information System Use. *University of California*, Graduate School of Library & Information Science, 120 Powell Library Building, 405 Hilgard Avenue, Los Angeles, CA 90024. Robert M. Hayes and Harold Borko. $196,640 for 24 months beginning August 14, 1979.

The purpose of this research is to examine four issues relating to evaluation of information system usage. Each issue is described in terms of a mathematical model which will be considered as hypothesis to be tested on data descriptive of various kinds of information systems. The first deals with usage as a function of allocation of resources, the second with distribution of usage, the third with significance of usage, and the fourth with distance and usage.

The project will list the hypotheses, validate the models, and analyze the consequences of them. Two phases are involved. In the first, the problems in data acquisition will be identified, methods for data acquisition will be pilot tested, and the underlying mathematics will be analyzed; in the second, the four hypotheses will be tested and the mathematical consequences will be explored toward the aim of developing design equations.

#018

Investigating Computational Gains Using Partial Information. *University of California*, Engineering Systems Department, 405 Hilgard Avenue, Los Angeles, CA 90024. Judea Pearl. $68,751 for 12 months beginning March 9, 1979.

The primary objective of this project is to develop a quantitative model describing the exchange between complexity and precision in non-numerical computer systems. This is one aspect of the general problem of representing information by "abstracts" which compensate for their incompleteness by their compactness, measured by the number of bits required to store the data in memory. This research addresses the general conditions under which a small imprecision can result in substantial computational savings, to quantify the laxity-complexity exchanges in practical computer systems.

#019

File Allocation and Query Optimization for Distributed Data Bases. *University of California*, School of Engineering and Applied Science, Hilgard Avenue, Los Angeles, CA 90024. Wesley W. Chu. $112,220 for 24 months beginning July 17, 1980.

The advent of computer network and microprocessor technology has established a technological foundation for implementing distributed computing and distributed databases on a computer network. Linking remote databases together not only provides more economical service but also allows for real time retrieval, update, and distribution of information.

Using a dynamic programming approach, based on query operation graphs, which provides for progressive elimination of undesirable strategies, the investigator proposes to design an algorithm which finds the optimal file allocation and query processing strategy and requires a relatively small amount of computation compared to the computation required to evaluate all possible strategies. This algorithm will consist of two phases: the first phase involves determining the values for parameters such as query frequencies, file sizes, communication costs, and processing costs; the second phase involves making the sequence of decisions that leads to the optimal strategy.

#020

Studies in the Organization of Information Structures. *University of California*, School of Engineering and Applied Science, Hilgard Avenue, Los Angeles, CA 90024. Judea Pearl and Norman C. Dalkey. $39,174 for 18 months beginning March 21, 1980.

It is proposed to study a statistical model of information systems, conceived in terms of probabilistic linkage between possible reports (data, messages, or other information) and events to be predicted. Typical information systems are large and complex, with much marginally relevant data. The proposed work is directed toward simplifying and modifying systems in such a way that decision-making becomes manageable and yet remains minimally degraded by the approximating transformations. Emphasis will be on procedures aimed at reducing the amount of information needed for decisions; extending capabilities for using partial or incomplete information; and clarifying the role of judgment in decision.

#021

Data Semantics and Logical Database Design. *University of California*, School of Engineering and Applied Science, Hilgard Avenue, Los Angeles, CA 90024. D. Stott Parker. $57,489 for 24 months beginning July 30, 1980.

With the increasing need for organized information systems, a great deal of attention has been focussed recently on the problems of database design. Typically a design process will proceed through three phases: (i) analysis and specification of user requirements, (ii) logical design of a database, and (iii) physical design for the logical database. The second of these, logical design, plays a key role in enabling the computer to meet user needs comprehensibly and efficiently in a changing information environment. Logical design of relational databases involves the development of a schema, or collection of tables, in which user data may be accurately stored. The measure of "accuracy" is the extent to which the schema reflects the data semantics, and permits the reconstruction, through joining tables together, of the original data.

This research concentrates on several topics currently of central importance in logical design: (1) data dependencies and normal forms, (2) design and implementation of relational views, (3) improvement or enhancement of the relational data model, and (4) development of effective methodologies for logical design. Project goals include unification of the data dependency field, resolution of several basic problems in the relational model (the use of null values in tables, the loss of semantic information connecting tables in the schema, and the ambiguity problems of multiple access paths between tables), and development of dependency-based schema design algorithms.

#022

An Information Processing Theory of Procedural Learning. *Carnegie-Mellon University*, Department of Psychology, Pittsburgh, PA 15213. Patrick W. Langley. $53,715 for 24 months beginning July 30, 1979.

The purpose of this research is to develop the framework for artificial systems which are capable of learning procedures. The development of such systems would preclude the present necessity for encoding large knowledge bases manually and thus make it economically feasible to produce information processing systems which would exhibit an "intelligence" not currently possible. Such systems would have the capability of automatically identifying the intrinsic knowledge content of large stores of information.

#023

Distributed Data Base Design. *Carnegie-Mellon University*, School of Urban and Public Affairs, 500 Forbes Avenue, Pittsburgh, PA 15213. William D. Haseman. $39,211 for 6 months beginning December 7, 1979.

Recent advances in computer technology, both hardware and software, have brought about a rapid change in how organizations collect, process and store data. Many of the large central computer facilities are giving way to clusters of smaller computers as rapid decreases in hardware prices eliminate the historical economies to scale advantages. The area which this proposal addresses is that of designing and distributing database structures over these networks which are good or "optimal" in some sense.

This project will develop a methodology and model for designing distributed databases. The model is initially viewed as a two-stage process involving first the design of the database and then the distribution of that database. The design of the database is accomplished using an integer program which minimizes the sum of access and storage cost. The distribution process uses a location model for determining the number of copies of each file and then their location. Initially these models will be solved sequentially in an interactive manner. Since this will only lead to a local optimum, it is then planned to use the knowledge gained to develop a single model which will combine both steps into one. Solution techniques for this model will be developed and both models will be tested using a medical database example.

#024

Theoretical and Experimental Investigation of the Dynamics of Information. *Case Western Reserve University*, Cleveland, OH 44106. William Goffman. $71,588 for 24 months beginning April 12, 1978.

The purpose of this research is to attempt to establish some of the fundamental principles which govern the behavior of communication processes and their associated information systems. It will have theoretical as well as practical implications. The theoretical portion of the research will deal with structural and dynamic properties of information transfer.

The goals of the structural aspect is to achieve a subject classification which is defined by the literature content rather than by the preconceived notions of a group of indexers and classifiers. This would be especially important for emerging disciplines and interdisciplinary fields where no classical structure exists. This project will use the mathematics of graph theory to establish the theoretical basis for imposing this type of structure on a population of information conveying elements such as journal articles.

The dynamic properties of communication among the members of a population such as scientists have been shown to behave with certain differences in accordance with an epidemic model such as governs the spread of infectious diseases. This project will examine these differences and develop an integrated model in which direct contact and intermediate host communication can take place simultaneously. The various relationships will be described in terms of differential equations.

The practical implications will derive from testing the theoretical findings experimentally on a sample of articles from the ***Annual Review of Information Science and Technology***. Information science qualifies as an emerging interdisciplinary area. In addition, its selection as the demonstration case will permit an examination of the field in terms of its future directions and boundaries.

#025

Language and Human Information Processing. *City University of New York*, Queens College, Department of Computer Science, 65-30 Kissena Boulevard, Flushing, NY 11367. John Moyne. $91,140 for 15 months beginning May 30, 1980.

This research is concerned with the communication processes which are involved in the transfer of knowledge from databases to users. An information bank or database is viewed as a body of knowledge which must be analyzed, segmented, classified and understood to be of any use. The instrument of communication which can specify what piece or pieces of information are needed from the database is called a language. The research is concerned with comprehension of the content of the database and the comprehension of the language. It is concerned with the organization and structure of the database and the techniques for the retrieval of the data only to the extent that they are relevant to the process of comprehension.

The investigator will examine issues in the theory of the nature, organization, and processing of information and the complexity of models of comprehension. He will survey the state-of-the-art and develop consistent formalizations of significant existing models in order to provide an apparatus to measure the potentials of such models and their claims to generality with regard to the psychological reality of comprehension and communication. The goal is to develop a universal family of specialized languages suitable for information processing.

#026

Analysis of the Operational Parameters for the Management of an Academic Library. *Clarkson College of Technology*, Potsdam, NY 13676. Michael Bommer and Ronald Chorba. $76,152 for 12 months beginning August 30, 1978.

This 12-month project aims to develop a set of operational parameters suitable for performance assessment and decision-making in academic libraries and information services. The study includes gathering base-line data on the kinds of, and manner in which, decisions are currently made in academic libraries.

Conclusions will be of value to managers of libraries and information centers, as well as to the administrators of the institutions in which they function.

The results, which will detail indicators essential to making resource allocation and acquisition decisions, will be particularly timely, considering the rising costs and declining budgets being faced by libraries and information centers.

Findings will be disseminated through library and information science journals, as well as through the science and engineering information literature.

#027

Designing Information Systems for More Effective Use in Decision Making. *University of Colorado*, Institute of Behavioral Science, Boulder, CO 80309. Kenneth R. Hammond, Richard L. Cook, and Robert Taylor. $105,531 for 12 months beginning May 2, 1979.

The hypothesis that cognitive aids in judgment and decision-making can improve the effectiveness of information systems is to be tested in this one-year project. The experiment will use two cognitive aids programs (POLICY and DIVERSE) linked to computerized data bases in a controlled manner that permits the searcher/user to cope with the complex task of human information and decision making processes. Results of this are to be compared with information searching and decision-making using (1) conventional manual methods, and (2) a computerized information system but without judgment and decision aids. Performance of the three groups is to be assessed on the basis of five criteria for evaluating effectiveness retraceability, reliability, interpersonal conflict reduction, and user satisfaction.

#028

Information Correlates of Basic Language Structures. *Columbia University*, Center for the Social Sciences, 420 West 118th Street, New York, NY 10027. Zellig S. Harris. $74,989 for 24 months beginning September 23, 1980.

This project is part of an ongoing effort to determine the informational structure of language. The research is aimed at reducing the information expressed by word-structure (morphology) to the already established information expressed by words combining into a sentence (syntax). Preliminary investigations suggest that this can be done, and have developed methods for doing it.

The whole work rests on a method of analyzing the structure and information of language which has shown that two processes—distinguished by objective and computable criteria—go into the construction of every sentence: one a direct informational processing of combining words, and the other a computable process of reducing low-information words which does not change the meaning of the sentence. The present project aims to extend this analysis to the internal structure of words.

Such an extension would show that morphology contains no information not contained in syntax. Hence it would show that the known process of combining free simple words into sentences carries all the information carryable by language. It would also give the starting point for computer programs to replace complex words by constructions of simple words, in any computer processing of written information.

#029

Normal Sentence Forms and Information Structure. *Columbia University*, Center for the Social Sciences, Box 20, Low Memorial Library, New York, NY 10027. Charles D. Parsons and Isaac Levi. $354,377 for 24 months beginning September 6, 1979.

This proposal will undertake a multi-faceted, interdisciplinary investigation of the structure of information within a framework provided by a theory of language which was developed by Professor Zellig Harris at the University of Pennsylvania. The research, which involves investigators from Columbia University and the University of Pennsylvania, will utilize the methods of distributional and combinatorial analysis, statistics, mathematical logic, and the philosophy of science.

The project will seek to determine what categories and relations of categories express information, how subsidiary information is connected to main information, how change of information is housed in information structures, how overt information can also carry covert information, and how information can be checked for its truth-conditions. The various investigations will be unified in that they will all start with the same representation of the way language carries information, and all will seek structures of information that go beyond the structure of language.

#030

Research on Combining Document and Data Retrieval Technologies. *Cornell University*, Department of Computer Science, Ithaca, NY 14850. Gerard Salton. $133,555 for 24 months beginning November 18, 1978.

There are currently three rather different approaches to the problem of information retrieval: first, the structured data base system used in business processing where a small number of attributes are used to achieve exact match retrieval; second, the reference retrieval or document processing systems where a large number of attributes are necessary and only partial matches can be achieved which result in lowered precision and increased recall; third, the emerging question-answering systems where the file is completely open-ended and new facts may need to be inferred, and reasoning required to generate appropriate responses to queries.

The purpose of this research is to combine the data and document retrieval technologies in order to achieve a single system capable of the exact matching required by business applications and the approximate matching needed in document retrieval and question-answering. Such a combined system would narrow the artificial separations which now exist between the different aspects of the retrieval problem, and create a more flexible environment, responsive to user needs.

#031

Mathematical Modelling in Automatic Information Retrieval. *Cornell University*, Department of Computer Science, Ithaca, NY 14853. Gerard Salton. $145,040 for 24 months beginning May 3, 1979.

The purpose of this research is to improve understanding of information retrieval processes, hence to improved retrieval systems. This proposal covers a number of tasks aimed at reconciling the results obtained by using different retrieval theories, and at explaining a number of open problems in the area. The aim is to derive a unified retrieval theory capable of representing classes of retrieval operations and applicable to many retrieval environments.

Two retrieval theories which will be investigated by this project are based on vector processing considerations and probabilistic methodologies respectively. The two theories produce apparently different prescriptions for computing the value of a term for indexing purposes. These differences will be examined and resolved if possible. Another class of questions involving relevance indications related to the terms matching characteristics between queries and records for purposes of interactive retrieval will also be investigated.

#032

Fast Computational Processes in Information Retrieval. *Cornell University*, Department of Computer Science, Ithaca, NY 14853. Gerard Salton. $29,605 for 12 months beginning July 27, 1979.

In recent years various approaches, including both hardware and software, have been introduced for use in information retrieval systems. Among the problems that have attracted special attention are those which are concerned with matching query statements with document content. In general, the efforts were directed toward achieving "perfect match" retrieval where records whose keyword assignment exactly matches a Boolean combination of terms included in a query.

This research will examine the efficiency and effectiveness of information retrieval applications using a hybrid system consisting of a fast computational array processor and a conventional general purpose computer. The associative processor is used for rapid computation and comparison of query statements with document contents for "closest match" as opposed to exact match retrieval. Within this system configuration, the computationally intensive tasks are executed on the fast, parallel floating point processor, data storage is done with the usual random access devices, while the conventional text processing operations and other control operations are executed on the general purpose computer. The project will also study the impacts of such systems on information processing cost and speed, as well as on the design of more flexible information retrieval systems.

#033

ANSI Committee Z39. *Council of National Library Associations*, 461 West Lancaster Avenue, Haverford, PA 19041. James L. Wood and Robert W. Frase. $20,000 for 12 months beginning May 10, 1979.

Committee Z39 of the American National Standards Institute (ANSI) develops standards relevant to information systems, library activities, document delivery, information dissemination, and information and data handling in all media. Based on findings made by a task force sponsored by the Foundation and the National Commission on Libraries and Information Science (NCLIS), the scope of Z39 was expanded, its program strengthened, and the management of Z39 reorganized. Since 1965, when the Foundation started supporting Z39, 28 standards have been developed by Z39. In addition, Z39, as ANSI's official representation to the International Organization for Standardization (ISO), has advanced the United States' interests in the information activities to ISO's decision-making meetings.

During the grant period, Z39 will start work on reaffirming, revising or withdrawing 8 standards, and on developing standards for 11 areas which have been selected by the information community.

#034

Conference on a National-Level Program of Access to Machine-Readable Data Files. *Data Use & Access Laboratories, Inc.*, Arlington, VA 22209. John C. Beresford. $68,300 for 12 months beginning December 12, 1977.

Publicly available statistics in the form of machine-readable data files contain much demographic, economic, and social information needed for the accomplishment of social science research and for a better understanding of the problems confronted by those who work in the public interest. Unfortunately, machine-readable data files are not routinely acquired, stored, catalogued, or processed within the existing library framework, and access to them is currently so complex that it has become a field of study itself.

One solution of growing interest is the establishment of a centralized national-level program to develop and implement standards for information services. In order to address the questions this raises and to sort out the many intellectual and technical tasks it would entail, a three-day conference will be held in the Spring of 1978. The conference will bring together representatives of some 30 major social science, information science, and documentation organizations and will include about 50 participants all told. The conference will be guided by an advisory committee and will be preceded by a small planning conference among key participants. Issues relating to the development of standard cataloging procedures, the operational components of a central clearinghouse, and the products and services that would result from a coordinated effort will constitute the core agenda. The conference proceedings will be widely announced and disseminated to professionals and potential users in both the research and the public administration communities.

#035

Natural Language Processing Systems and Ill-Formed Input. *University of Delaware*, Department of Computer and Information Sciences, Newark, DE 19711. Ralph M. Weischedel. $136,681 for 24 months beginning August 4, 1980.

Although natural language interfaces to databases have improved markedly in recent years, they have only begun to consider ill-formed input which is one of the major obstacles to truly natural language systems. Quite often natural language queries will be ill-formed in the sense of being misspelled, ungrammatical, or not entirely meaningful. A requirement for any successful natural language interface must be that the system either guesses at a user's intent, requests direct clarification, or at the very least, identifies the nature of the ill-formedness in an intelligent way. In this project the researchers will investigate the hypothesis that ill-formedness is rule-based. Violation of the rules of normal processing will be used to signal ill-formedness. Meta-rules that modify the rules of normal processing will then be used for error identification and recovery.

It is anticipated that this approach will ease system development and processing. First, because it will be possible to design the normative processing system independent of the error identification and recovery methods. Second, not invoking ill-formedness processing until normal processing fails, avoids unnecessary run-time costs for well-formed sentences which are the usual type of input. And third, describing ill-formedness through meta-rules that relate to normative rules will avoid duplication of aspects of normative processing and allow general statements covering classes of ill-formedness.

This project will be performed in collaboration with Norman K. Sondheimer, whose participation will be supported by Sperry Univac, Blue Bell, Pennsylvania.

#036

The Design and Assessment of Individualized Instruction in Bibliographic Data Base Access. *Drexel University*, Graduate School of Library Science, Philadelphia, PA 19104. Charles T. Meadow. $229,600 for 24 months beginning February 13, 1978.

The Drexel University Graduate School of Library Science and the Franklin Institute Research Laboratories, are continuing, in this joint effort, initiated under SIS76-09737, to design and evaluate a system providing individualized instruction in accessing and retrieving computer-stored information files. The evaluated computer-based aids expected as products of their research should provide a wide range of information users with alternatives to querying systems through information intermediaries and/or to learning to use systems in training sessions not designed for their needs.

Products of the 12-month initial phase include a working model of the instruction system, and initial evaluation of its feasibility, cost/effectiveness, and transportability to a variety of computer environments. This phase will concentrate on further design and assessment of the sub-component of the system devoted to diagnostic and assistance programs which monitor and provide feedback to users attempting to improve their search abilities. In addition to a more amplified working model of the systems, second phase results are expected to provide input into the design of future storage and retrieval systems. The long-range goal to which this project will contribute is the availability of systems with built-in capacity to assist users in forming their information requests, and in locating and retrieving information relevant to their needs.

#037

Factors Influencing Network Sharing of Computer Resources in Research. *EDUCOM*, P. O. Box 364, Princeton, NJ 08540. James C. Emery. $362,269 for 20 months beginning September 28, 1978.

Several organized experiments in resource sharing by computer networks are underway. This project will study the operation of ten resource sharing organizations to analyze the impact of their services on the use of network information services by research scientists. The study will include regional computer networks, discipline-based networks, and special services such as systems for bibliographic search and retrieval. Site visits will be made to the organizations, and interviews will be conducted with managers and with users. Survey data will be collected and analyzed.

The objectives of this research are: (1) to develop a clearer understanding of resource-sharing and the way it is influenced by characteristics of the resources, the users, and the communication processes; (2) to determine what effect computer resource-sharing has on the efficiency and effectiveness of research using it; and (3) to provide guidance for organizations concerned with resource-sharing at the policy-making and operational levels.

#038

Electronic Information Exchange in Microcomputer Component Standardization Research. *Electronic Industries Association*, 2001 Eye Street, NW, Washington, DC 20006. John F. Hessman. $77,113 for 18 months beginning July 24, 1978.

The Joint Electronic Device Engineering Council (JEDEC), a standing committee of the Electronic Industries Association, will explore the use of electronic information exchange for selected aspects of its work on hardware and software standardization of microcomputer/large scale integration (LSI) products. Via the NSF-sponsored electronic information exchange (EIES), JEDEC will structure and moderate standardization discussions among industry representatives drawn from regular standardization committees and task groups of JEDEC. Observers of this process, also via EIES, will include representatives of such other interested groups as computer scientists, computer hobbyists, professional societies, trade associations, the U.S. government (military), and the trade press. Electronic information exchange will be compared with conventional means of conducting JEDEC standardization activities with respect to: amount and effectiveness of communication, elapsed time in processing standardization decisions, ability to resolve standardization issues, and cost-effectiveness in financial and other respects.

#039

Development of Theory and Techniques for Automated Knowledge Utilization, Transfer and Extension. *University of Florida*, Gainesville, FL 32611. Julius T. Tou. $166,740 for 24 months beginning August 3, 1979.

This research project is aimed at exploring the design concepts and techniques for automated knowledge utilization, transfer and extension. During the initial stages, efforts will be concentrated on the development of theories and methods for the representation and organization of knowledge and for knowledge-based design and growth. This investigation will be followed by the study of the requirements and procedures for on-line knowledge transfer, interactive browsing, and selection techniques for knowledge transfer. The total project is slated for two years and is expected to contribute to our understanding of the fundamentals of knowledge processing.

#040

Degradation of Accuracy in Social Network Data. *University of Florida*, Gainesville, FL 32611. H. Russell Bernard. $19,403 for 12 months beginning December 18, 1979.

Scientists desire to improve the methodology of information science research, particularly the methodology for identifying and studying information networks. They also desire to develop measures of reliability and validity for network data. This project will assess the nature and degree of human error in data gathered by sociometric techniques from information networks. It will also assess the quality of the data that are collected by providing a comparison of sociometric survey data with electronically recorded logs of social interactions. The research will be carried out with the cooperation of volunteers from research groups now using the electronic information exchange system at the New Jersey Institute of Technology. Started at the West Virginia University, this research will be completed at the University of Florida where the Principal Investigator is now located.

#041

Multiple Attribute Retrieval Techniques for Tree-Structured Information. *Florida State University*, School of Arts and Sciences, Tallahassee, FL 32306. Caroline M. Eastman. $24,063 for 12 months beginning August 7, 1980.

Search algorithms based upon simultaneous use of several attributes of the data will be investigated. Such "multikey" searches are used in many types of information retrieval systems, including both document retrieval systems and database management systems.

The various multikey search algorithms, functioning and proposed, can be classified by the type(s) of query handled, by the evaluation metric for responses and on the basis of the underlying file structure or database. The focus of this research is on tree-structured algorithms. Internal nodes of a tree contain information to guide the search by indicating how well the records in their respective subtrees match the query.

Analytic approaches have typically made restrictive assumptions while experimentation has been unwieldy. The proposed work seeks a different course, posing a complex model of a class of search algorithms and analyzing its performance with artificially generated distributions (involving so-called K-d trees) as well as some complex real data.

#042

A Study of the Potential Influence of Social, Economic, Regulatory and Technological Factors on Scientific and Technical Communication Through 2000 A.D. *Forecasting International, Ltd.*, 1001 North Highland Street, Arlington, VA 22201. Norman Nisenoff. $178,775 for 24 months beginning September 13, 1978.

The purpose of the project is to improve our understanding of the interaction of technological and non-technological factors influencing future changes in the scientific and technical information (STI) transfer process. This research will examine the influence of social, economic, regulatory and technological events upon established trends in the processing and distribution of STI, and predict the consequences of such trends and possible future events. The trends studied will related to both capabilities and needs within predicted environments for the year 1980, 1990 and 2000, corresponding to the occurrence or non-occurrence of selected subsets of events.

Descriptions of alternative futures will be developed by the application of trend analysis and cross-impact analysis using existing data bases as a nucleus. This is to be augmented by individual as well as collective judgment during a workshop session involving expert consultants.

#043

Support for Coordination of Scientific and Technical Information Activities. *Franklin Institute Research Laboratories*, Philadelphia, PA 19103. Alec Peters. $50,000 for 7 months beginning December 4, 1979.

The purpose of this project is to provide technical, consultative, analytic and study support services to the Foundation's Division of Information Science and Technology (IST) in its conduct of coordination activities at the national and international levels.

The contractor shall perform the following tasks as shall be further specified by IST: (1) Logistical and administrative arrangements for conferences on topics to be designated by IST; and (2) Preparation of analytical papers on designated topics which focus on special issues relevant for coordination of national and international information activities.

This continues work started under NSF Award DSI77-14816.

#044

The Effects of STI Changes on the Engineering Profession. *The Futures Group*, 124 Hebron Avenue, Glastonbury, CT 06033. Hedvah L. Shuchman. $129,927 for 15 months beginning June 13, 1978.

This study is designed to examine the effects of changes in scientific and technical information transfer on the engineering profession. The objectives of the study: are to identify the information used by technologists of various disciplines and specialties; identify the major modes of communication and patterns of use; describe changes which may be anticipated in information use patterns and communication modes as a result of changes in information storage and dissemination methods; and identify policy implications resulting from these evolutionary changes.

By means of a large scale survey a broad spectrum of engineers will be asked to provide information on their use of S&T information. Data collected will be analyzed to provide information profile and use patterns of engineers, and to project changes of the profile and pattern resulting from advances in information processing technologies. This research, which will cover technologists in industry, government and academic research, will be of value to information service managers and policy makers.

#045

Syntactic Structure of Information and Information Processes. *Georgia Institute of Technology*, Atlanta, GA 30331. Vladimir Slamecka. $137,790 for 24 months beginning August 1, 1979.

This research is intended to increase fundamental understanding of the syntactic structure of information and information processes. The purpose is to gain a better understanding of human processing of sign shapes (words) as measured by information transfer rates and to determine the redundancy curve for natural language. This research will rely on the ediometer, an instrument developed at the Georgia Institute of Technology for measuring the deviance of artificial words from the norm.

#046

Study of Potential New Applications of Information Theory to Problems of Information Science. *Georgia Institute of Technology*, School of Information and Computer Science, Atlanta, GA 30332. Pranas Zunde. $82,449 for 12 months beginning September 10, 1979.

The purpose of this study is to: (1) critically evaluate what impact information theory has had on the broad domain of information science; (2) identify problem areas of information science to which information theory is potentially applicable; and (3) explore promising extensions of theory towards broader-based foundations of information science. This is to be achieved by an in-depth study of the first two items by the investigating team, followed by a workshop of a small group of recognized experts who will concentrate their deliberations principally on the third item.

#047

Verification of Syllable Hypotheses in Automatic Phonetic Analysis of Spoken English. *Haskins Laboratories, Inc.*, New Haven, CT 06510. Franklin S. Cooper. $109,475 for 12 months beginning August 30, 1979.

Automatic phonetic analysis of fluent speech is an essential component of computer speech-processing systems. Dealing successfully with fluent speech requires that the system make use of knowledge about the language (top-down processing) as well as parametric representation of the speech waveforms (bottom-up processing). At an intermediate level, the parametric representation is segmented and categorized into a sequence of symbols and words, resulting in a drastic reduction in the information rate. This research is focussed on the problem of parameter-to-symbol transformations.

#048

Research on Automated Natural Language Generation. *Hofstra University*, Comparative Literature and Languages, 100 Fulton Avenue, Hempstead, NY 11550. Karen J. Neff. $110,633 for 24 months beginning August 7, 1980.

Recent years have witnessed considerable research devoted to natural language processing by computer. This progress has been substantiated. But this work has tended to concentrate on machine analysis, understanding, and inference patterns and has largely ignored the question of how to produce a variety of intelligent outputs. There are no known computer programs that will accept an input topic and freely generate a basic informative paragraph on that topic.

This project will: (a) describe in formal terms the structured properties of basic non-narrative paragraphs; (b) investigate the connections between the syntactic form of such paragraphs and their semantic content, or the type of information they convey; (c) examine the proper use of stylistic variation in sentences and paragraphs by introducing such variation into a paragraph-generating computer program; and (d) show how paragraph patterns of information in text can be algorithmically identified.

#049

Video Full Text Processing. *IIT Research Institute*, 10 West 35th Street, Chicago, IL 60616. Peter B. Schipma. $222,570 for 24 months beginning August 23, 1978.

Problems of using video technology for full-text delivery are being explored through simulation of a video text system in which the output will be displayed on a standard television screen. This research project will consider: (1) methods for overcoming the low resolving power of the standard television screen through electronic reformatting techniques; (2) human factors analysis of a video text system; (3) trade-offs of efficiency among electronic text distribution, hybrid systems, and major alternatives. These factors will be studied using available standard television and computer equipment to manipulate video images of text and to display them to selected user groups. Reading efficiency performance parameters for a high controlled user's group will be studied, using the shared task paradigm to identify optimal system design parameters. Later, using the optimal parameters, use statistics will be studied for several subject populations to whom the system will be made available.

#050

An Augmented English Dictionary for Information Retrieval. *IIT Research Institute,* 10 West 35th Street, Chicago, IL 60616. David S. Becker. $83,285 for 12 months beginning June 26, 1980.

This proposal addresses problems of information retrieval systems, including ambiguous words, incomplete or inconsistent indexing, and terminology not recognized or understood by the searcher and others.

The project aims to study and develop algorithms and computer programs which automatically disambiguate and index terms, thereby providing an augmented English dictionary for use in information retrieval. Specific objectives include: study and develop methods for disambiguating and indexing terms, based on keyword methods, conditional probability, and methods from artificial intelligence, test and measure the reliability of such approaches; characterize the structure of the dictionary and the nature and import of its operations and failure modes if any.

#051

Impact of a Paperless Society on the Research Library of the Future. *University of Illinois,* Library Research Center, Urbana, IL 61801. F. W. Lancaster. $50,540 for 12 months beginning June 11, 1978.

The purpose of the project is to investigate the implications of paperless communication systems for the research library. The study essentially involves two facets: (1) a projection of the impact of present and future technology on the production and distribution of primary and secondary literature; and (2) an understanding of the changing roles and functions of research libraries in this evolutionary environment. Several alternative scenarios of library activities up to the year 2000 will be depicted assuming differing levels of automation. Policy and operation issues to be explored and investigated include possible interrelationships among future libraries and between libraries and publishers, the potential roles of the librarians and information specialists, the requirements for the library profession and library services, and other problems affecting the various areas of the library.

#052

Theoretical Issues in Natural Language Processing-2: An Interdisciplinary Conference. *University of Illinois,* Coordinated Science Laboratory, Urbana, IL 61801. David L. Waltz. $14,242 for 6 months beginning June 26, 1978.

This award is to support an interdisciplinary conference to be held July 25-27 at the University of Illinois at Urbana-Champaign. It will bring together researchers interested in using computer-languages to represent and process the texts of documents prepared in natural languages. Results will be applicable to the design and development of advanced information storage and retrieval systems which (1) allow utilization without specialized knowledge of computer languages; and (2) provide more assistance to users in accessing and retrieving required information.

The conference is expected to attract about three hundred researchers in linguistics, computer science, information science, and cognitive and social psychology. The conference will have six sessions each consisting of 5-6 invited papers from outstanding researchers along with one or two papers selected from those submitted to the conference.

The collection of papers to be produced will be widely circulated and is expected to familiarize both interested researchers and others with current ideas and paradigms in storing and processing natural language text.

#053

Storage and Access Costs in Data Structures for Information Retrieval Problems. *University of Illinois,* Coordinated Science Laboratory, 1101 West Springfield, Urbana, IL 61801. Donna J. Brown. $62,663 for 24 months beginning August 14, 1980.

It is proposed to study efficient storage and accessing of data in information retrieval systems. Two objectives of information retrieval systems are (i) to store data as compactly as possible and (ii) to be able to quickly retrieve information and perform desired updates. There are, however, tradeoffs between the efficient storage and accessing of data, and these tradeoffs are dependent on the precise application. One phase of the proposed work is to study particular data structures and the storage-access cost tradeoffs inherent in performing certain operations. A second phase of the work will be development of analysis techniques which can be applied to various retrieval problems, using cost measures appropriate to the particular applications.

Preliminary work has been directed toward obtaining lower bounds, indicating inherent limitations in storage and retrieval of information. What is particularly encouraging is that these lower bounds are essentially tight for the implementation schemes considered. Thus, determining lower bounds has led to "optimal" retrieval methods. It is proposed to continue this work, considering more complicated data representation schemes and more complicated cost measures.

#054

Modeling of Human Behavior in Seeking and Generating Information. *University of Illinois,* Coordinated Science Laboratory, College of Engineering, Urbana, IL 61801. William B. Rouse. $144,120 for 24 months beginning December 21, 1979.

The goal of this research is to develop analytical techniques that are applicable to the design of computer systems that understand their users and can adapt the human-computer interface so as to best meet the needs of the particular human using the system at the moment.

Development of mathematical models of human behavior in information generation and retrieval will be undertaken. The specific tasks to be studied include text editing and searching/browsing through data bases. The models will be based on Markov process theory, and will be formed by analyzing the human's keyboard entries. The resulting models will be useful for various applications, especially in the area of training.

#055

An Integrated Man/Machine Interface to Facilitate Network Resource Utilization. *University of Illinois,* Urbana, IL 61801. Martha E. Williams. $174,432 for 18 months beginning February 14, 1978.

The purpose of this research is to investigate the complexity of factors which are barriers to the use of machine-readable files of the scientific and technical literature. The majority of the world's currently published abstracting and indexing literature is in computer-readable form. Unfortunately there are no standards governing the format of data bases, the elements and vocabularies included in them, or the on-line systems for searching them. The lack of standards imposes requirements on searchers to become familiar with the variety of existing data bases, vocabularies, access protocols, system features and command languages as well as to keep up-to-date with changes in all of them.

Many of these differences could be made less apparent by developing translators for converting the procedures or conventions or terminology of one system into equivalent procedures, conventions, and terminology of other systems. These conventors or translators which make systems appear to be alike to the searcher can be called transparency aids. An integrated collection of transparency aids then would comprise a transparent system. This project will: (1) design a generalized model for a transparent system, (2) analyze the alternative factors affecting the transparency interface, (3) develop a skeletal user-oriented terminal for accessing the transparency system, and (4) integrate the transparency subsystems. The alternatives study will show advantages and disadvantages of various model configurations, and the user-oriented terminal development together with the subsystem integration will permit evaluation of the various subsystems by showing how well they interact to perform their functions.

#056

Automatic Duplicate Detection of Journal Articles Appearing in Multiple Databases—A Feasibility Study. *University of Illinois,* Information Retrieval Research Laboratory, 291 Coordinated Science Laboratory, Urbana, IL 61801. Martha E. Williams. $269,303 for 24 months beginning February 13, 1980.

The purpose of the study is to develop a methodology for automatically detecting overlap among database records found in multiple databases. There are currently some 71 million bibliographic records in computer-readable form. The problem of overlap among them represents duplicative efforts in database generation; it represents duplicative efforts in storing and processing records for researchers who search multiple databases for a particular query. Thus, duplication represents a real cost to producers, processors, and users of databases.

The project will develop a methodology for automatically determining overlap among databases. It will test and implement this methodology on seven databases: AGRICOLA, BIOSIS, Previews, CACon, Excerpta Medica, Medline, Science Citation Index (SCI), and Social Science Citation Index (SSCI). The proposed methodology involves analysis of the databases to determine common data elements in order to construct effective keys for matching records in multiple files. The study will produce overlap statistics among the seven databases, data element frequency statistics, and a methodology for duplicate detection that might be used by database producers or database vendors. It could serve as a preprocessing step to reduce storage and manipulation costs or as a post retrieval step to eliminate the duplicative output delivered to users.

#057

Access Time Versus Storage Space in Information Retrieval Systems. *University of Illinois,* Urbana, IL 61801. Swomitri Swamy. $59,829 for 24 months beginning June 26, 1980.

It is proposed to study the performance of algorithms that retrieve information from a direct access file system. Specifically, the aim of the work is to derive results on the tradeoff between access time and storage space.

Because there are many types of query (e.g., exact match, partial match, range query) and many types of algorithms (based on, e.g., probability theory, queuing theory, combinatorics), it is unrealistic to expect to find a theory universally applicable to these disparate search problems and algorithms. Therefore to render the problem manageable the class of retrieval systems will be narrowed (i) by restricting the kinds of information that is stored in the records and (ii) by considering searching problems where the query concern some mathematical domain. The first restriction will constrain the kinds of answers that a system may provide, while the second will make possible use of the mathematical relationships in achieving the goal of relating parameters of storage space and access time.

The results will take the form of nontrivial lower and upper bounds for access time as a function of storage space. The upper bounds will involve the construction of specific retrieval schemes that achieve those bounds.

#058

Structuring the Lexicon and the Thesaurus with Lexical-Semantic Relations. *Illinois Institute of Technology,* Computer Science Department, 3300 South Federal Street, Chicago, IL 60616. Martha W. Evens. $43,575 for 24 months beginning August 6, 1979.

The development of hierarchical structures, such as indexes and abstracts, to enable ease of access to large stores of information is a major concern of information science. The advent of computerized information processing systems has made it possible to generate some of these tools, such as indexes, automatically. But the computer also offers the potential for much more powerful information retrieval such as question-answering systems and other models of natural language processing. If this potential is to be realized, more powerful processing tools which express more and deeper relationships among words than synonymy and autonomy must be made explicit for the machine. The question is: Can these lexical-semantic relations be identified automatically?

The first step in this research will be to identify new lexical-semantic relationships and to derive their logical properties. The second step will be to extract word-relation-word triples from dictionaries. The target will be the expanded system of relations developed in Step 1. The third step is to design a thesaurus using these relations and conduct a pilot study of its effectiveness in enhancing recall.

#059

The Economics of Library Subscriptions: A Sample Survey. *Indiana University,* Bloomington, IN 47401. Herbert S. White. $61,025 for 12 months beginning June 13, 1978.

The objective of this project is to examine the impact of periodical availability in libraries on the placement and cancellation of subscriptions. Decreasing subscriptions by individuals, the impact of inflation on library budgets, and the growth of networks and consortia are believed by many to have seriously eroded publisher income. The reseachers will develop information and data on the role played by each of the above factors on subscription levels. The results will lead to a better understanding of the magnitude of today's problems, will assist managers of information services in their planning, and will provide policy makers with hard data and information on which to base their recommendations.

#060

Problems in Decision-Making and Information Retrieval Studied with Techniques of Fuzzy Sets and Systems. *Indiana University,* Department of Mathematics, P. O. Box F, Bloomington, IN 47401. Madan L. Puri. $112,690 for 36 months beginning December 5, 1979.

The objective of this project is to develop and use techniques and results of fuzzy sets and systems theory to provide realistic solutions of problems in (1) decision-making under uncertainty; (2) presence of imprecision in information retrieval systems; and (3) subjective evaluation of random phenomena in information retrieval systems. Models of information retrieval systems incorporating the concept of fuzzy sets and systems will be developed which can be used to facilitate accurate decisions even if the input information (data) is only approximately known. In many areas of information science, such as pattern recognition, decision-making, information retrieval, and management information systems, among others, human judgment is often imprecise and the decision-maker manipulates not with numbers but with vague concepts. These vague concepts are technically known as "fuzzy sets." By studying information retrieval systems with a continuum of information, new models for large information systems can be built which will take into account approximate grouping of the information items and the possibility of measuring the subjective information contained in these groups. Eight specific decision-making and information retrieval problems and their solutions in terms of the concept of fuzzy sets and systems are outlined as specific research tasks.

#061

Electronic Information Exchange in Research on Devices for the Disabled. *Innovative Systems Research, Inc.*, 103 Cooper River Plaza East, Pennsauken, NJ 08109. Jane McCarroll. $20,275 for 12 months beginning June 6, 1979.

The objective of this project is to conduct a planning meeting and pilot test of the use of electronic information exchange in the development of an improved nationwide delivery system for rehabilitation technology products, services, and information.

The project has been funded, as an addition to DSI77-17924, by the Division of Information Science and Technology with funds provided by the Foundation's Directorate for Applied Science and Research Applications (ASRA), as part of the program for "Science and Technology to Aid the Physically Handicapped."

#062

The Structure and Development of the Information Science Literature: A Citation Study. *Institute for Scientific Information*, Philadelphia, PA 19101. Henry G. Small and Gabriel Pinski. $32,689 for 12 months beginning August 23, 1978.

A bibliometric and citation analysis of the field of information science will be carried out which will utilize citation and publication data from the *Social Sciences Citation Index* (*SSCI*). A nine-year (1969-1977) file of data will be extracted from the SSCI and mounted on an on-line system for manipulation and graphic display of citation networks. This on-line system will permit analyses of the internal structure and development of the field of information science.

In the second phase of the 12-month study, a second data base, a clustering version of a three year (1972-1974) SSCI file, will be used to investigate connections between information science and other areas, and to conduct a comparative study of information science and the field of education research.

#063

Research Needs and Applications for Indicators Based on the Scientific and Technical Literature. *Institute for Scientific Information*, 325 Chestnut Street, Philadelphia, PA 19106. H. Roberts Coward. $18,142 for 6 months beginning September 11, 1978.

This award is for a two-conference project which will bring together information scientists, analytical methodologists, and science studies researchers for the purpose of: (1) assessing the existing body of research concerning quantitative aspects of the scientific and technical literature; (2) identifying in this existing research indicators with applications in studies of the social system of science and science policy; and (3) identifying areas of information science research that would enhance applications in research on the scientific enterprise and science policy by improving such indicators.

#064

User Requirements for Scientific Reviews. *Institute for Scientific Information*, 325 Chestnut Street, Philadelphia, PA 19106. Susan E. Cozzens. $31,421 for 12 months beginning November 7, 1978.

Increasing the use of scientific reviews has been proposed as a solution to the problems of the "Information Explosion," because reviews evaluate and synthesize research results and present them in a condensed form. One of the important functions served by review articles is for current awareness outside of one's own speciality. Therefore, the efficient use of reviews serves to cross-fertilize ideas and innovations in science.

This study will examine the relationship between users in a variety of science-related roles and the characteristics of the reviews they prefer to use. It is posited that level of research activity will be related to the frequency of review-use for current awareness, and that the level of codification of the researcher's field will affect the extent to which reviews are used for bibliographic retrieval. The kind of review desired for current awareness will vary, in turn, by the frequency of use. Frequent users are expected to need more detailed, authoritative, and current reviews.

A questionnaire mailed to 1200 scientists and engineers will provide the primary survey instrument.

#065

Citation-Based Measures of Interactivity and Similarity. *Institute for Scientific Information*, 325 Chestnut Street, Philadelphia, PA 19106. Elliot Noma. $38,188 for 12 months beginning March 6, 1980.

Bibliographic citation patterns contain important information on the interactions among publishing entities. In order to extract this information from the various kinds of publishing data, it is necessary to formulate appropriate measures and associated statistical methods. In previous and ongoing work the principal investigator has developed a measure of the flow of influence among publishing entities. An important property of this measure is that it is not distorted by variations in size or by variations in the amount of self-citation.

The proposed work involves the development of measures for other important aspects of bibliographic systems, specifically the intensity of interaction and the similarity of two entities with respect to how each relates to various third parties. The similarity measure will provide the basis for creation of multidimensional maps using the technology of multidimensional scaling. These maps will in turn be compared, using a cluster similarity measure, newly developed by the principal investigator. This latter measure provides a tool for further exploration and validation of the interaction and similarity measures mentioned above.

The methodology to be developed will be applicable in a variety of ways. For example, the cluster similarity measure would make it possible to trace the historical evolution of the bibliographic relationships in a discipline. More generally, the statistics of cluster resemblance would enhance the usefulness of multidimensional scaling for its widespread use in bibliometric and psychometric studies. Finally, interaction and similarity are relevant not only to bibliometric studies but also to indexing and classification.

#066

Human Requirements in Information Retrieval. *Johns Hopkins University*, Baltimore, MD 21218. Alphonse Chapanis and Gerald D. Weeks. $195,841 for 24 months beginning February 5, 1979.

The purpose of this research is to examine the human requirements in the use of information systems. The principal objectives are to: (a) develop, refine and apply research and evaluation methodologies appropriate to making these requirements explicit; (b) collect quantitative data on these user-system behaviors; and (c) present the finds in a form that will be useful to designers in the development of more effective information retrieval systems.

The project involves three major tasks. The first will obtain: (a) a realistic estimate of recall and precision ratios of documents in different areas of science; (b) an estimate of the inter-rater reliability of expert judgments of document relevance in selected areas of science, and (c) a better appreciation of how users assess the "relevancy" of documents and the cognitive processes they use in those assessments.

The second task will obtain quantitative data on a number of aspects of man-system interactions and construct a model of an information retrieval system with quantitative statements of its human requirements.

The third task is a controlled experiment comparing the effectiveness of query formulation developed between users and information specialists who communicate in three modes: interactive teletype, telephone, or face-to-face.

#067

Improving the Use of Computer-Based Information Resources Within the Engineering Classroom. *University of Kansas*, Center for Research, Inc., 2291 Irving Hill Drive, Lawrence, KS 66045. Dale Rummer. $61,421 for 15 months beginning January 1, 1978.

This award continues research initiated in 1976 under DSI76-08949 to test and evaluate the costs and usefulness of using computer-based information resources as a teaching tool in a range of engineering courses. Materials for teaching access to and use of information services have been developed and introduced into engineering design courses at the University of Kansas. Utilization of these materials by students, the cost of using these information systems to teach the courses and the effects on students and faculty have been monitored and assessed. In addition to the report summarizing results of the first year, a workshop was held at the 1977 ASEE Summer School to inform chemical engineering faculty about the progress of the study.

The second year effort will revise materials and strategies developed in light of feedback; expand the universe of students; and evaluate longer-term effects on participants in 1976-77 courses. Transportable teaching materials, guidelines and strategies for their use, and assessments of the costs and effects of using these materials will be the products of the second year effort. A workshop is planned for summer 1978, as one means of disseminating the results of the project.

#068

Development of Statistical Indicators of Scientific and Technical Communication. *King Research, Inc.*, 6000 Executive Boulevard, Rockville, MD 20852. Donald W. King. $9,281, November 15, 1977.

Costs adjustment to a predecessor award described below.

#069

Data Gathering for Statistical Indicators of Scientific and Technical Communication from Primary Sources. *King Research, Inc.*, 6000 Executive Boulevard, Rockville, MD 20852. Donald W. King. $147,442 for 12 months beginning November 22, 1977.

This is the third year of a continuing effort to develop statistical indicators of scientific and technical communication. This third-year effort will result in the development of the first detailed model that will allow us to trace the flow and predict the impact of changes that are likely to occur in scientific and technical communication as a result of changes in, for example, copyright legislation, library automation and page charge policy. This model will also specifically provide data and analyses on the impact of Federal funding on scientific and technical communication.

#070

Systems Analysis of Scientific and Technical Communication in the United States. *King Research, Inc.*, 6000 Executive Boulevard, Rockville, MD 20852. Donald W. King. $11,394 for 6 months beginning December 8, 1977.

The number of scientific and technical articles published doubles every eight to ten years. The range of possible uses for the information in those articles and the number of users are also increasing rapidly. Therefore the limits of what can be communicated by printing, mailing, storing, and retrieving pieces of paper may be at hand. One possibility for improving communication is to use computer systems to capture and transmit new information. This is not the only alternative, but it is one which holds great promise for the productive use of information. The objective of this project is to devise an analytic framework which will allow publishers and other information distributors to judge the potential of the electronic alternative to paper-based communication in relation to the conventional system. This framework is intended for use in comparing total communication systems in terms of costs, benefits, and other implications defining issues; and in specifying the data needed for the resolution of those issues.

This amendment to Contract NSF-C76-15515 provides for consideration of the impact of the revised copyright law on the analysis and on the cost model developed as part of the project.

#071

Evaluation of Information Science Research. *King Research, Inc.*, 6000 Executive Boulevard, Rockville, MD 20852. Donald W. King. $30,000 for 6 months beginning September 30, 1978.

This six-month award provides for the development of a design to carry out an evaluation of the information science research program. The approach includes selecting representative information science research projects and developing methods for assessing their contribution to information science and technology. Products of the award will include: (1) selection of the projects to be evaluated; (2) development of measures of use and utilization of research results; and (3) design of instruments and analysis procedures for implementation of the design.

Several design studies, in addition to this one, are being conducted concurrently. Phase II awards will implement the most promising evaluation designs produced.

#072

Evaluation Measures and Methods for On-Line Bibliographic Systems. *King Research, Inc.*, 6000 Executive Boulevard, Rockville, MD 20852. Nancy K. Roderer and Robert R.V. Wiederkehr. $143,846 for 15 months beginning August 14, 1979.

On-line bibliographic searching is an important relatively recent development in information services. A range of evaluation measures can be identified for on-line bibliographic systems. Some are based on measures previously developed for manual systems and batch processed systems, but others relate to the interactive capabilities available on line. The purpose of this project is to review existing evaluation measures and to develop new ones for on-line bibliographic systems. Both information science literature and material in related fields such as computer systems, data base management, and numeric data bases will be reviewed.

The evaluation will involve the application of cost-benefit analysis to the overall bibliographic system, with parallel evaluation at function, product and service, and activity levels. Evaluation will be viewed from the perspective of all the major bibliographic system participants, including abstracting and indexing (A&I) services, information dissemination centers, libraries, information brokers, and users. Both A&I and library cataloging data bases will be considered.

Study tasks include development of an on-line system model, analysis of evaluation measures and methods, and selected testing of new approaches. Output of the study will be a monograph explicating evaluation results.

#073

Uses of Citing Statements to Improve Retrieval. *Lehigh University*, Bethlehem, PA 18015. John J. O'Connor. $116,037 for 24 months beginning July 5, 1978.

Computer-readable full text of scientific journals is becoming increasingly available as computer typesetting is more widely used for economic reasons. When papers are available in computer-readable full text they can be retrieved by text searching. Such papers can also be employed to provide useful computer-readable information about other papers, which are not necessarily themselves in computer-readable form.

Specifically, citing statements within papers represent what users have found important in the papers cited. If citing statements can be computer-identified in computer-readable full text, they can then be used to: (1) improve the reference retrieval performance of present retrieval services, (2) facilitate user processing of retrieval outputs, and (3) increase the usefulness of citation indexes.

This project will investigate procedures for citing statement recognition, and for the use of computer-recognized citing statements to improve reference retrieval and retrieval output processing.

#074

A Theoretical Basis for Knowledge Management. *Lehigh University,* Center for Information & Computer Science, Building 8, Bethlehem, PA 18015. Donald J. Hillman. $123,551 for 24 months beginning June 19, 1980.

The purpose of this research is to provide a theoretical basis for the development of intelligent systems particularly those that involve the management of knowledge and inference. Such systems will require that the capabilities possessed by existing information processing systems be substantially expanded.

The project will construct a first-order theory that will relate the common features of knowledge management systems to a set of underlying concepts and deductive rules. It will then examine the completeness of the theory and investigate its implications for system design. The first-order predicate calculus will be employed, and a relational approach to empirical data will be taken.

#075

Citing Statements: Recognition by Computer and Use to Improve Full-Text Retrieval. *Lehigh University,* Center for Information and Computer Science, Mart Library #8, Bethlehem, PA 18015. John J. O'Connor. $115,447 for 24 months beginning July 22, 1980.

Computer-readable full text of scientific journals is becoming increasingly available as computer typesetting is more widely used for economic reasons. Large-scale, full-text retrieval services for lawyers have existed for a number of years. An advantage of such retrieval is that the full text of a publication contains far more information about itself than do its abstract and index terms. It also contains information about other publications, i.e., those it cites. If citing statements can be automatically recognized, they could provide very powerful access tools that could not only identify related documents but also link related concepts within documents.

The investigator will study the following questions: (1) Can procedures for computer recognition of citing statements developed for one subject field (inorganic chemistry) be used effectively in other subject areas?; (2) Can the use of citing statements aid retrieval if full text is searched rather than index terms and abstract text?; (3) How much false retrieval will be produced by the use of citing statements to aid retrieval?

#076

Evaluation of the Use of Slow-Scan Television and Telefacsimile in a Library Network. *Library of Congress,* Federal Library Committee, Navy Yard Annex, Room 400, Washington, DC 20540. James P. Riley and Joan M. Maier. $145,905 for 24 months beginning August 23, 1978.

In a cooperative effort, fifteen Federal and state libraries will be experimenting with the use of slow-scan television and telefacsimile devices to transmit images of document pages over telephone lines for the purpose of testing the concept and practicability of electronic document delivery in a nationwide network environment. Impact on library patrons, librarians, and library administrators will be assessed. Data on network activities will be collected and analyzed by an independent contractor. In addition to interlibrary loans, the system will also be tested for its usefulness for teleconferencing, work sessions, training, and other appropriate applications.

#077

Representation and Processing of Spatially Structured Information in Cognition. *University of Louisville,* Systems Science Institute, Louisville, KY 40208. Bruce A. Whitehead. $77,840 for 24 months beginning September 5, 1980.

Knowledge about space forms a basic framework for processing information about the world. The principles of representing and processing such spatial knowledge are the topic of the proposed research. The objectives are to better understand these principles, both within the specific constraints of human information processing and in the more general case in which some of these constraints are relaxed. The feasibility of such principles under well-defined constraints will be investigated and tested in a simulation model. Cognitive operators for spatially structured information will be modeled on the basis of both physiological and psychological data. The modeling technique will bring this human data to bear upon principles stated in the broader framework of information science. The generalizability of these representation and processing principles will then be explored.

#078

A Simultaneous Equation Model of the Demand for Academic and Professional Journals. *University of Lowell,* Department of Economics, One University Avenue, Lowell, MA 01854. Carol C. McDonough. $25,342 for 12 months beginning May 9, 1980.

The objective of this proposal is to construct and estimate a simultaneous equation model of the demand for academic and professional journals.

A four-equation model will be developed that describes both individual and institutional journal demand, journal cost, journal price, and journal product (measured as number of pages) in terms of parameters that are exogenous to the model. Such variables include: journal quality, size of the potential journal market, average annual journal costs, number of journals published by the publisher, and the technical/scientific nature of a journal, and a journal's subsidy/profit relationship with its publisher.

#079

Research on Behavioral Variables Related to Information Systems. *University of Maryland,* College of Library and Information Services, College Park, MD 20742. Jerry S. Kidd. $163,011 for 18 months beginning August 22, 1979.

The project will utilize techniques borrowed from cognitive psychology and artificial intelligence (scripts/frames) in order to identify a scientist's context and status of research at any particular point in a research project. The purpose is to be able to tailor information products to meet the specific requirements of the research process as they change during the various phases of a project.

The major task will be to determine whether it will be possible to prepare realistic, detailed scripts and frames.

The study is intended as the first step toward the ultimate goal of developing an information system which will interact with a scientist in diagnosing his/her information needs, in constructing a schema which places these needs in the context of what he already knows, and in matching this knowledge structure (schema) with scientific information contained in the system data bases.

#080

Applications of Logic and Theorem Proving to Databases. *University of Maryland,* Department of Computer Science, College Park, MD 20742. Jack Minker and Barry E. Jacobs. $207,835 for 36 months beginning April 18, 1980.

The major thrust of this research is the application of automated theorem proving and mathematical logic to database investigations. Three types will be treated: (1) theorem proving by linear resolution with unrestricted selection function using trees (LUST theorem proving), (2) first-order logic and databases, and (3) database logic. Database logic is a recently developed formalism whereby existing database systems can be viewed in terms of logic. It has the prospect of providing a unifying theoretical framework for database systems.

This project continues research supported under NSF Grant MCS73-03433 A02.

#081

A Generalized Model of Document Retrieval. *University of Massachusetts,* Department of Computer and Information Science, Amherst, MA 01003. W. Bruce Croft. $61,740 for 24 months beginning June 27, 1980.

The purpose of this research is to enhance both the understanding of information retrieval processes and the capabilities for designing more effective retrieval systems. The project includes tasks to develop a probabilistic retrieval model and to evaluate performance of the model by applying several measures to two test collections of documents, queries and relevance judgments.

Specifically, the research will extend the probabilistic models based on binary index terms (those of Spark-Jones & Robertson and Van Rijsbergen) to include weighted index terms. Two conditional probabilities are introduced consisting first of the probability that a term occurs in a relevant document, and then the probability that the term occurs in a given document description. The retrieval results will be evaluated using two test collections and applying recall and precision tables of E values, and a measure relating total number of relevant documents retrieved to queries.

#082

Experimental Evaluation of Information Delivered in Various Modes and Combination of Modes. *Massachusetts Institute of Technology,* Center for Advanced Engineering Study, Cambridge, MA 02139. Karen C. Cohen. $131,746 for 24 months beginning September 26, 1979.

The purpose of this study is to evaluate the effectiveness of various modes of information transfer. Modes to be assessed will include: lecture, reprints in print form, facsimile reprints accessed on-line, and audio cassettes of lectures. Effectiveness will be defined and examined in terms of user characteristics, user criteria for selection of one or more modes, and actual transfer of knowledge.

The study will access the various modes singly and in combinations. The critical events will be lectures delivered at annual and national meetings of the American Institute of Chemical Engineers. The alternative modes will be reproductions of the contents of these lectures.

Results should interest information science researchers concerned with understanding the effects of modes on transfer of information, and systems designers who wish to improve the delivery of factual information.

#083

The Use of Intonation in Automatic Speech Understanding. *Massachusetts Institute of Technology,* Center for Cognitive Science, Cambridge, MA 02139. Janet Pierrehumbert. $43,701 for 24 months beginning July 3, 1980.

Intonation, as expressed in the fundamental frequency of the voice is related to the location of stress in words and to syntactic phrasing. The proposed research is directed towards using the intonation system to help a computer system, for understanding speech, determine what words it has heard, and how these words are organized into sentences. It is anticipated that work in this area will contribute to making speech understanding systems good enough to be useful in many practical applications, and that the research will also add to the understanding of how people process speech. Pilot work suggests that people use a number of qualitatively different intonation patterns even when asked to speak neutrally. Because of this variety, the location of stressed syllables and phrase boundaries cannot be identified from the fundamental frequency contour. The investigator proposes to design a computer program which would screen possible analyses of an utterance, rejecting an analysis if the observed fundamental fequency contour represented an impossible intonation for the stress and phrasing hypothesized.

#084

A Research Program on Electronic Document-Delivery Systems for Interlibrary Loans. *Massachusetts Institute of Technology*, Electronic Systems Laboratory, 77 Massachusetts Avenue, Cambridge, MA 02139. J. Francis Reintjes. $83,725 for 12 months beginning February 14, 1979.

The project sets forth a two-year research program pertaining to electronic document delivery systems applicable to interlibrary loans. In the first year, a study will be made of technological alternatives for the various segments of an electronic document delivery network, including information encoding methods, data compression schemes, communications channels and receiving end facilities. Non-technological issues that impact upon the technological organization of the network will also be studied. In the second year, experimental evaluations will be made of selected basic approaches that show promise of contributing substantially to electronic interlibrary loan networks. An effort will also be made to establish an experimental two-mode network as a means of assessing the viability of the electronic approach from technological and non-technological viewpoints.

#085

Electronic Information Exchange in Mental Workload Research. *Massachusetts Institute of Technology*, Cambridge, MA 02139. Thomas B. Sheridan. $148,454 for 18 months beginning July 19, 1978.

The problem of defining and measuring mental workload is a critical one in complex man-machine systems such as aircraft piloting and traffic control, nuclear reactors, and various military systems where there are large risks to public safety and large concentrations of capital under the control of one or few persons. In such systems perceptual and decision demands may change rapidly from a state of boredom when automation functions perfectly to a state of high stress and operator unreliability when automation fails.

Thirty researchers and engineers concerned with mental workload in the USA, Canada, and Great Britain will employ the New Jersey Institute of Technology's facility for research on electronic information exchange to define terms and experimental questions; share knowledge of applicable theory and research; organize and critique research in progress; and draft, edit, review and publish archival papers. These papers are to be transferred to files publicly accessible by anyone through a commercial computer network, and the availability and use of this services is to be advertised throughout the broader community of interest.

#086

Investigations of Computer-Aided Document Search Strategies. *Massachusetts Institute of Technology*, Laboratory for Information and Decision Systems, Cambridge, MA 02139. Richard S. Marcus and J. Francis Reintjes. $155,360 for 24 months beginning June 11, 1980.

The purpose of this research is to simplify the tasks which confront users of online retrieval systems in their attempts to gain access to desired information. The difficulty of learning and using existing systems is demonstrated by the fact that those who ultimately need and use the information derived from the retrieval systems by and large cannot and do not operate the systems themselves.

This project is aimed at investigating how the user's effectiveness and efficiency of searching can be improved by providing him with computerized assistance in the formulation and execution of his search strategy. More specifically it will develop and test procedures and dynamic aids that will help to optimize search strategy formulation in conjunction with initiatives taken by the user. These techniques derive from an emerging theory of online search optimization which calls for a combination of factors including a knowledge base, computational linguistics, and computer-assisted instruction techniques to automatically select appropriate search strategies and data bases.

#087

Investigations of Electronic Interlibrary Resource-Sharing Networks. *Massachusetts Institute of Technology*, Laboratory for Information and Decision Systems, 77 Massachusetts Avenue, Cambridge, MA 02139. J. Francis Reintjes. $95,000 for 12 months beginning June 11, 1980.

This is a continuation of NSF Grant IST78-23251 entitled "A Research Program on Electronic Document-Delivery Systems for Interlibrary Loans".

#088

Natural Computation and Control. *Massachusetts Institute of Technology*, Cambridge, MA 02139. Whitman A. Richards. $234,196 for 12 months beginning April 11, 1980.

Research in natural computation explores the information processing mechanisms utilized by biological systems in performing basic sensory and motor functions. The MIT effort is primarily concerned with understanding the computational mechanisms which underlie the human visual system's representations of the external world. Topics to be investigated will include: (1) stereopsis, (2) texture, (3) structure from motion, (4) transparency, (5) color perception and normalization, and (6) derivation of object shapes from contour and shading.

This project continues research supported by NSF Grant MCS77-07569.

#089

Computer-Based System for Sharing Technical Information. *University of Michigan*, Center for Research on Learning and Teaching, Ann Arbor, MI 48109. Karl L. Zinn. $20,000 for 12 months beginning May 5, 1978.

This project is an extension of a research grant testing the feasibility of using teleconferencing and electronic-publishing techniques to facilitate the sharing of scientific and technical information among individuals in a scientific community. The information to be shared is stored in a generally accessible computer system. A computer/communication network allows one who originates information to disseminate it to an audience of potential users without the delay, expense or editorial constraint of conventional publication. The continuation will carry the project activity further into the communities involved, and examine the effects in greater depth through interviews. Most participants are from universities, others are from government and industry.

#090

Computer Studies in Formal Linguistics. *University of Michigan*, Department of Computer and Communication Sciences, 2076 Frieze Building, Ann Arbor, MI 48109. Joyce Friedman. $87,508 for 24 months beginning June 28, 1979.

The purpose of the proposed project is to study logic-based formal systems for natural language, using the computer as a research tool. Formal linguistic theories can explicate and clarify the relationships among syntax, semantics, and pragmatics; computer implementation imposes a processing view which adds a new dimension to the theory. Montague grammar provides a well-defined fragment in which important questions can be phrased precisely and answers can be formulated and tested.

The main goals of the project are (1) to use mathematical and computational techniques to explore the syntactic and logic components of Montague grammar, (2) to develop further and more formally the notions of dynamic interpretations, and (3) to investigate their application to problems in pragmatics.

The results of the project will be LISP programs and formal papers. These will be usable by other linguists in doing research and in teaching formal grammar.

#091

Links of Interest and Expertise Among Scientists. *University of Michigan*, Ann Arbor, MI 48109. Manfred Kochen. $49,621 for 12 months beginning August 14, 1978.

Connections among scientists based on interests and expertise are vital for the scientific enterprise in the selection of reviewers and in building reputations through recommendations. Scientists use the connection network to amplify their individual capacities to cope with the volume and difficulty of accumulating scientific knowledge. Understanding the network of connections is necessary for providing information systems with a scientific foundation and for clarifying the issues regarding the healthy growth of scientific knowledge. The researchers will examine the functioning of coupling networks and their impact on organizing specialty clusters of scientists; develop and test new methods for observing how specialty clusters are formed; and develop a method of describing interests and competence to contribute and to evaluate research.

#092

Dynamics in the Organization of Scientific Specialties. *University of Michigan*, Mental Health Research Institute, Ann Arbor, MI 48109. Manfred Kochen. $134,782 for 24 months beginning August 7, 1980.

The processes whereby scientists add to knowledge in their specialties affect the structure of scientific information. Understanding how specialties grow helps in designing information processing systems for scientists. Communication among specialists and the organization of knowledge play a key role.

Four tasks will be undertaken to explicate the growth of specialties: (1) search for a representation of a specialty state; (2) derivation of difference equations for how the state-variables change; (3) analysis of these nonlinear difference equations; (4) evaluation and extension of this model to deal with many specialties simultaneously and to characterize knowledge spaces and the invariance of laws expressing knowledge dynamics under transformation from one representation to another.

These tasks will be accomplished primarily by mathematical analysis, construction and use of computer programs to generate and compare representations and sensitivity analysis of the difference equations. Some empirical studies, using samples of mathematical literature to trace the use of theorems, methods, conjectures and concepts in the generation of new knowledge units will be undertaken to justify the dynamic model and to compare the implied direction of change in the state-variables with past trends.

#093

A Conference on Serial Collections in a Network Environment. *Minnesota Higher Education Coordinating Board*, St. Paul, MN 55101. Alice E. Wilcox. $32,469 for 12 months beginning September 29, 1978.

This conference will bring together 6 leading researchers and 30 academic librarians to suggest guidelines for the development and managment of serials collections in academic libraries participating in resource-sharing networks. The attendees will examine and evaluate the results of a number of research efforts performed by the 6 researchers and will synthesize these results into practical guidelines for use by academic librarians in their planning and management functions.

#094

Systematic Process Evaluation of Fundamental Information Science Research. *Mitre Corporation*, Metrek Division, 1820 Dolley Madison Boulevard, McLean, VA 22101. Karen Levitan. $29,847 for 6 months beginning September 26, 1978.

This six-month award provides for the development of a design to carry out an evaluation of the information science program. A systems approach to evaluation is taken in the study. It will attempt to identify and describe the set of interrelated variables, which influence program outcomes. Its products will include: mapping of the field of information science; a design for examining non-NSF sources of funding for information science research; an initial description of the current state of knowledge in information science as well as related fields, and the characteristics of proposed information science research.

Several design studies, in addition to this one, are being conducted concurrently. Phase II awards will implement the most promising evaluation designs produced.

#095

International Scientific and Technical Information Programs. *National Academy of Sciences*, 2102 Constitution Avenue, NW, Washington, DC 20418. Judith A. Werdel. $96,500 for 12 months beginning March 6, 1979.

This award provides support for continuation of U.S. participation in the International Federation for Documentation (FID) and the International Council of Scientific Unions/Abstracting Board (ICSU/AB), as well as for continuation of certain information clearinghouse services within the National Academy of Sciences-National Research Council (NAS-NRC) that relate to international scientific and technical (S&T) information. Support for 1979 U.S. dues to both international organizations is included in the award, as is support for the U.S. National Committee for FID and ICSU/AB. FID and ICSU/AB are two of the principal international organizations concerned with S&T information and information technology. Through them, the U.S. scientific, technical and information communities cooperate internationally in many programs directed toward development of S&T documentation and the organization exchange, and dissemination of information for S&T abstracting and indexing services.

Clearinghouse and information services within the NAS-NRC that are supported by this award constitute an important U.S. resource for the collection and referral of information on international, bilateral, and foreign national S&T information organizations, programs, and issues.

#096

Refinement and Use of Concept Similarities in Information Retrieval. *National Bureau of Standards*, Institute for Computer Sciences and Technology, Building 225, Washington, DC 20234. Matthew B. Koll. $35,498 for 12 months beginning September 1, 1980.

It is well established in information retrieval theory that the ability of a system to retrieve relevant documents is dependent on the agreement between the system's and users' concepts. In designing IR systems, this is often stated as a vector space model of concept similarities. This model is the focus of the proposed research. The aim of this study is to maximize the agreement between the concept similarities derived from text, and those perceived similarities held in the population of data base users. The project will also examine methods by which IR systems can use the enhanced concept relations. The final output will be the design of an IR system which utilizes such enhancements.

The methodology involves human users as well as experimental machine runs. Users' views about the similarity of concept pairs will be obtained from questionnaires. System's views will be obtained from term co-occurrence data, weightings, and word distance measures. By measuring the closeness between these two views, predictions will be made regarding the use of the statistical techniques as a means to simulate the existing semantic relations. The research will test various algorithms for creating the concept space from term co-occurrence data in the text of a data base.

#097

Support of the Numerical Data Advisory Board (NDAB). *National Bureau of Standards*, Office of Standard Reference Data, Room A323, Physics Building, Washington, DC 20234. James E. Loy. $12,300 for 12 months beginning August 30, 1979.

The Numerical Data Advisory Board (NDAB) of the National Academy of Sciences (NAS) provides a national forum for the consideration of the whole range of problems associated with the compilation and dissemination of scientific data. Through the establishment of working groups of carefully selected experts the NDAB carries out detailed studies and submits reports relating to the generation, processing, storage, retrieval and use of critically evaluated data. Among the efforts planned for the coming year are two of special interest to the Foundation: "encouraging the development and application of new methods of data management," and "identifying specific areas of data management and information science in need of further research and development."

The National Bureau of Standards is the executive agent for Federal agencies contributing to support of the NDAB program.

#098

Electronic Information Exchange in Planning for White House Conference on Library and Information Services. *National Commission on Libraries and Information Science*, 1717 K Street, NW, Washington, DC 20036. Marilyn Gell. $40,335 for 7 months beginning August 27, 1979.

In November 1979, the National Commission on Libraries and Information Science will hold the first White House Conference on Library and Information Services. The conference will be attended by more than 1100 delegates, drawn from the information professions and the general public. Its goal is to examine a variety of ways in which libraries and information centers can more effectively bring information to the people of this country. To facilitate planning for the Conference, the Commission will make use of the Electronic Information Exchange System, a research facility developed and operated by the New Jersey Institute of Technology under contract with the Foundation. Users of this facility can exchange messages, participate in conferences, and collaborate in the preparation of documents even though they are dispersed throughout the country and utilize the system on different schedules. The Foundation has already made the facility's resources available to twenty users to be named by the Commission. This award allows the Commission to designate eighteen additional users and to acquire the further resources needed for effective employment of the facility by all thirty-eight of its users.

#099

Study of Impacts of a Formal Computer-Based Information System on Information Networks. *Naval Ocean Systems Center*, System Development Branch, San Diego, CA 92152. Benjamin A. Saltzer. $71,891 for 18 months beginning May 23, 1978.

The project is to determine and quantify the impacts of formal computer-based information systems on the flow of scientific and technical information within and among engineering groups. These may include changes in the quantity and quality of the flow, the key roles, and changes in the informal intergroup information network. A balanced mix of quantitative and qualitative techniques is to be used to gather and analyze data collected "before," "during," and "after" the introduction of computerized information systems to the groups. Participants in the study are groups of engineers and scientists working in the area of sonar and undersea acoustic systems.

#100

Human Attention and Performance Limitations in Processing Visual Information. *University of Nebraska*, 209 Burnett Hall, Lincoln, NE 68588. John H. Flowers. $34,672 for 12 months beginning July 30, 1979.

The research is concerned with the ability of human observers to allocate attention to spatial regions of a visual display, and their ability to ignore irrelevant visual information which appears in the vicinity of their attentional "focus." One series of experiments will examine how structural properties and familiarity presented in one portion of a brief visual display affect accuracy for recalling or recognizing material appearing in other regions of the display (e.g., attention "capture"). A second sequence of experiments will study interference effects caused by distractors which appear in the spatial vicinity of a fixated "target," and will examine how the focus of the attentional "beam" can be influenced by a number of task variables which change the state of the observer. The data obtained will be used in the evaluation and development of models of attention which will provide clues for the design of artificial systems capable of selective abstracting of input information.

#101

Electronic Information Exchange Test Facility. *New Jersey Institute of Technology*, 323 High Street, Newark, NJ 07102. Murray Turoff. $275,197 for 18 months beginning August 3, 1978.

The Electronic Information Exchange System (EIES) is operated as a facility for experimental research on the use of computer-based communication and information resources by geographically dispersed individuals. Based on a dedicated mini-computer with more than 200 million characters of disc storage available for direct support of users, the EIES offers such features as private message exchange, computer-based conferences, on-line notebooks, and an all-electronic journal. By means of TELENET, a commercial packet-switching telecommunications system, the full range of these capabilities can be made available to 300 participants in NSF research projects.

#102

Electronic Information Exchange Test Facility. *New Jersey Institute of Technology*, 323 High Street, Newark, NJ 07102. Murray Turoff. $2,000 for 0 months beginning March 28, 1980.

Supplement to IST-C77-21008.

#103

Probabilistic Measurement Theoretic Models of Information Processing. *New York University*, Washington Square, New York, NY 10003. Jean-Claude Falmagne. $156,909 for 31 months beginning April 2, 1980.

The goal of this project is to further mathematical understanding of a broad class of mechanisms that appear to arise in human information processing. The investigator will study probabilistic models intended to describe the transduction and recoding of input information into random variables, and how choices and decisions are made by comparison of the relevant random variables.

#104

An In-Depth Analysis of a Comprehensive Data Base on the Communication of Scientific/Scholarly Knowledge. *New York University*, Department of Economics, 538 Tisch Hall, Washington Square, New York, NY 10003. Fritz Machlup. $160,178 for 12 months beginning July 10, 1978.

This project is for additional and in-depth analyses of the data base developed by the principal investigator in the course of earlier research on the production and distribution of scientific and technological information. Data regarding the publication of books and journals will be analyzed in greater detail than previously possible in such areas as book sales, by field, over a period of 8-10 years, and sales data, by type of purchaser; costs of journal distribution, and ratios of operating costs to total costs. Subscription prices and circulation data will also be examined in order to estimate price elasticities of demand by type of subscribers. Additional analyses of library data will also be performed.

#105

Information Science: An Analysis of Methodological Issues and Interdisciplinary Relationships. *New York University*, Department of Economics, Washington Square, New York, NY 10003. Fritz Machlup. $39,973 for 12 months beginning May 27, 1980.

The purpose of this project is to analyze the various methodological issues and interdisciplinary relationships which arise from the fact that information science is not yet a fully integrated assemblage of systematic studies. It consists of various approaches addressed to investigations of the operations, devices, technology and organization, by which humans transmit and receive, and thereby transfer knowledge among one another. It is necessary to investigate the ways in which these various approaches interact in order to form a picture of the relationships involved.

Specialists in the following disciplines will be invited to prepare papers on the intra and interdisciplinary relationships as they see them: Library Science; Computer Science; Human Communications including Cognitive Psychology; Systems Theory; Communication Theory; Living Systems; Artificial Intelligence; Cybernetics; and Decision Theory. The Principal Investigator will analyze the logical and methodological processes in each of these approaches and synthesize a coherent description of the present status of information science.

#106

Publication of Proceedings of a Conference on Telecommunications and Productivity. *New York University*, Center for Science & Technology Policy, New York, NY 10003. Mitchell L. Moss. $9,976 for 12 months beginning April 8, 1980.

The Center for Science and Technology Policy at New York University held a two-day conference to explore the social and economic consequences of new telecommunications systems.

This conference identified emerging uses and trends and examined organizational factors affecting the application of telecommunications systems. The major issues discussed are: (1) Present and future uses of telecommunications systems in the home and office; (2) The potential impact of communications technologies on productivity in public and private organizations; and (3) Alternative policies for future research and development in telecommunications.

The proceedings of the conference will include: conference papers, summaries of the workshop sessions, and an interpretative summary of the complete conference. The proceedings will be published by the Addison-Wesley Publishing Company.

#107

Research into Methods for Automatic Classification and Fact Retrieval in Science Subfields. *New York University, Department of Linguistics,* New York, NY 10003. Naomi Sager. $211,000 for 24 months beginning December 29, 1977.

The purpose of this research is to expand the fundamental understandings which are essential to the eventual design of information systems which can analyze, process, combine and deliver desired information from natural language data bases. Such a capability would supplement existing document retrieval systems by operating on the natural language texts of the documents in order to determine whether desired factual relations are present in the data base. In some cases, this type of capability could even compare and compile data from different source documents by way of treating the natural language texts of the documents.

The first objective is to develop procedures and computer programs for obtaining the informational word classes of a science subfield and for classifying textwords according to these categories. The second objective is to extend our knowledge of information structures in science and to the compatibility of such structures with texts in the publishing literature. Thus, the research will develop automatic methods for generating word classes for a given subject matter and for tagging the individual words with the appropriate class markers. The development of such procedures are essential if advanced language-based, as opposed to document-based information processing is to be applied in diverse areas of science.

#108

Data Base Management Systems for Natural Language Data Bases. *New York University,* 15 Washington Place, New York, NY 10003. Naomi Sager. $188,943 for 24 months beginning September 7, 1978.

This project will explore the application of current data-base management systems (DBMS) to the management of natural language data bases. Natural language information structures will be compared with the information structures handled by existing DBMS of the hierarchical, network and relational language types. A prototype query language appropriate to natural-language information systems will be designed and evaluated. The results of the research should provide a clear picture of the technical requirements for extending data base technology for use with natural language data bases in the future.

#109

Computer-Based Investigation into the Structure of Information. *New York University,* Linguistics Department, New York NY 10003 Naomi Sager. $274,645 for 24 months beginning January 25, 1980

The proposed research will utilize information formats as the means for comparing and quantifying structural properties of information in different subject areas and on different levels of generality of information. An information format is a row-column matrix derived by formal and computerized procedures to sample texts in a given area such as a subfield of science into which successive sentences of texts in the field can be mapped. These information formats then contain the same information as the original texts.

It should be possible to obtain precise descriptions of three levels of information: (1) individual facts (as in reports); (2) factual summaries and generalization (as in survey articles); and (3) fact correlations and analysis (as in research articles). A hypothesis to be tested is that relations between the structures on the different levels can be established, showing how particular factual information is transformed to become the object of generalization and the operand of theoretical constructs.

#110

Computer Applications to Analytical Chemistry. *University of North Carolina,* Chapel Hill, NC 27514. Thomas L. Isenhour. $71,000 for 12 months beginning May 15, 1978.

Dr. Isenhour is studying further and developing the following computer applications to analytical chemistry.

A. Rank and factor analysis methods are being developed for routine use in determining the number of components in a mixture and isolating pure component spectra for file searching. Mixtures of Xylenes and saturated hydrocarbons are being used as test samples and the methods developed are being further tested on complex pesticide mixtures.

B. Infrared spectroscopy coupled with gas chromatography is being developed as a method of analysis of complex samples. Data collected in the Fourier domain are used to reconstruct the gas chromatogram much more efficiently than in the standard approach of transforming each individual interferogram. Further study is in progress, coupled with Part A, to determine background information, isolate pure spectra and develop automated procedures to analyze complex mixtures.

C. As data bases have become larger and better, a great deal of attention has been given to developing search systems. There is, at present, no standard method of comparing the capability of search systems. In this study an error propagation method of evaluating the performance of library search similarity measures is being developed and applied to various search systems of mass spectral data. The technique is based on repeated searches several nearest matches to selected compounds. Performance is being judged by the rate at which the dissimilarity between target compounds and propagated matches increases.

This research is a continuation of that currently supported by Foundation Grant CHE77-09762.

#111

Computer Applications to Analytical Chemistry. *University of North Carolina,* Chapel Hill, NC 27514. Thomas L. Isenhour. $66,500 for 12 months beginning June 11, 1980.

This research is a continuation of work currently supported by NSF and described above.

#112

Development of a Conceptual Framework to Guide Research in Information Science. *NTL Institute for Applied Behavioral Science,* 1501 Wilson Boulevard, Arlington, VA 22209. Edwin E. Olson and Dale Lake. $49,611 for 12 months beginning September 7, 1978.

This project aims to identify promising models and paradigms in the social and behavioral sciences; evaluate their application to information science problems; and encourage the information science community to use those methods, assessed to be useful, in their research.

A three-step process will carry out these objectives. First, a select group of accomplished social and behavioral scientists will be convened to develop an inventory of models with the potential to guide information science research. This will be followed by a symposium in which a larger group of scientists will present and discuss technical papers to critique and expand the output of the first session.

Results of the symposium will be made widely available to the information science research community, as well as to behavioral and social scientists.

#113

Message Spaces and Computing Time in Performance Allocation Processes. *Northwestern University,* Center for Mathematical Studies in Economics and Management, 633 Clark Street, Evanston, IL 60201. Kenneth R. Mount. $78,409 for 24 months beginning August 7, 1980.

Information plays both an implicit and explicit role in the allocation of economic products and factors of production. Firms and governments are large complex entities explicitly processing information about current and future technology, resources, demography, etc. Despite this important role in economic matters, information aspects have been studied largely as special topics, rather than as an integral component of the allocation process.

An investigation of informational and computational requirements for economic allocation mechanisms is proposed. Particular attention will be given to incorporating into the model the information processing costs internal to a particular economic unit. Internal costs are important because if they are ignored, then the merging of two units appears to decrease the information requirements of the system, whereas in fact merger merely internalizes communication between the merged units. Tradeoffs between the amount of information contained in economic systems and complexity of computation will be examined.

#114

Research on the Impact of a Computerized Circulation System on the Performance of a Large College Library. *Oberlin College,* Oberlin, OH 44074. Virginia B. Harris. $139,980 for 24 months beginning September 21, 1978.

The project is an in-depth study of the impact of an automated circulation system on the performance of a large college library. It will examine the relationships among three types of variables: (1) user attitudes, (2) system states (i.e., the operating status of the system), and (3) objective performance (i.e., the quality and quantity of the services rendered). Data collection will be integrated into the automated circulation system. Analysis and use of data will be integrated into regular library operations and management. One of the goals of the project is to develop and refine objective performance measurement mechanisms for library operations.

#115

The Probabilistic Matching and Control of Author Names in Automated Library Systems. *OCLC, Incorporated,* 1125 Kinnear Road, Columbus, OH 43212. Thomas B. Hickey. $42,321 for 12 months beginning August 21, 1979.

The process of author name matching is one of the most difficult and complex problems which confronts computerized bibliographic data bases such as library catalogs. The difficulties arise from such factors as: size, diversity of different author names, a high proportion of foreign names which have been romanized, and pseudonyms and other name variants. When names cannot be effectively matched against search queries the result can be "lost"records. In time these problems can distort the accuracy of a data base. The purpose of this research is to increase the role of computers in name-matching. This will make it possible for computers to create and monitor authority files with less human effort and greater data base integrity and searching ability.

The techniques and algorithms which will be investigated will enable a computerized system to automatically: correct and/or flag typographical errors in names; identify names in a data base that are similar to a name entered by a search query; measure the similarity among names by computing the probability that two or more names refer to the same person. These techniques will generalize to a wide variety of problems ranging from those encountered by libraries to those which are peculiar to abstracting and indexing services. It is expected, in particular, that the resultant techniques will form the basis for a new approach to authority control within automated library systems.

#116

Computer Terminal Requirements for On-line Catalogs in Libraries. *OCLC, Inc.*, 1125 Kinnear Road, Columbus, OH 43212. Neal K. Kaske. $97,902 for 12 months beginning May 22, 1980.

This proposal addresses problems related to the role of the library as both an agent of information transfer and a user of information technology. Specific project objectives include: develop measures of library use with particular emphasis on card catalog and online catalog use; identify patterns of human interactions with the two catalogs; develop and test an algorithm for estimating the number of public computer terminals needed by a library to support an online public catalog. Library use data will be collected at the Ohio State University Libraries which have both online and manual catalog systems.

This research will produce guidelines that libraries can use to estimate the number of computer terminals needed to support an online public catalog. These guidelines will be independent of library type and size, and will benefit individual libraries, systems of libraries, and library networks in their economic and systems planning.

#117

Extension and Applications of a Theory of Information Flow and Analysis: The Use of Information in a Decision-Making Environment. *Ohio State University*, Department of Computer Science, 1314 Kinnear Road, Columbus, OH 43210. Marshall C. Yovits. $105,670 for 24 months beginning July 16, 1979.

The purpose of this research is to develop a basic theory of information flow and analysis. It builds on previous research which has been underway at Ohio State University over the past several years. The approach relates information to its effectiveness and thus its use and value. Theory, measures, and relationships thus derived relate to the way information, once it is transmitted and received, is utilized by a decision-maker.

More specifically the research attempts to: (1) identify and quantify important variables and parameters in the information flow process; (2) establish relationships among these variables; (3) apply the theory to practical situations and to examine the resulting implications; and (4) develop models both simulation and experimental, to utilize and validate the theory.

#118

Concurrency in the Access and Manipulation of Information Structures. *University of Oregon*, Eugene, OR 97403. Carla Schlatter Ellis. $62,500 for 24 months beginning June 18, 1980.

One of the objectives of data base design is to collect the information needed for various related applications and organize it into a unified system that allows efficient access and modification. With many users sharing the data structures of a large data base, purely sequential operations may provide less than satisfactory performance.

Concurrent operations, in contrast, potentially improve performance but introduce new problems (e.g., indefinite postponement) and aggravate old ones (e.g., consistency). Thus an important area of investigation deals with techniques for increasing concurrency of operations (primarily search, insertion, and deletion) on the kinds of data structures employed in data base systems.

In sequential programming "good" solutions are often based on recognized techniques or data base structures. For parallel processing, more work is needed to build up a repertoire of techniques and data structures, to evaluate their performance, and to develop guidelines for their use. This background motivates the following proposed lines of research: (1) to identify and add to the set of techniques for *concurrent* manipulation of structured data, (2) to characterize the correctness issues associated with the application of these techniques, (3) to suggest a set of relevant performance measures, and (4) to provide a simulation tool for comparing solutions.

#119

Electronic Information Exchange in Legislative Technical Research. *Participation Systems, Inc.*, 43 Myrtle Terrace, Winchester, MA 01890. Chandler Harrison Stevens. $150,575 for 18 months beginning July 24, 1978.

This project is a trial of electronic exchange of information within the community of legislative technical researchers. This community took form in the early 1970's due to three intersecting tendencies: (1) development of state legislative staffs in general, (2) development of interests among scientists and engineers in public policy, and (3) development of various non-electronic networking techniques particularly suited to sharing technical information about legislative issues.

Electronic information exchange will be used in an interstate exchange of both substantive and process information. Initial emphasis will be given to exchanging technical inquiries from legislators and the responses obtained from scientists and engineers. In these inquiries formatted response briefs will be used. The format will be revised on-line during the course of the project, resulting in a nationally acceptable convention for the preparation of response briefs.

The body of data created during the project will be organized using the electronic medium. This new body of literature will be a valuable asset to legislative researchers which should continue to grow once the project ends.

#120

Correlation of Language Structure with Information. *University of Pennsylvania,* Philadelphia, PA 19104. Zellig S. Harris. $163,329 for 24 months beginning April 10, 1978.

The purpose of this research is to determine the structure of language-borne information. It will first attempt to establish classes of words which have special information characteristics. It will then assemble these words into combinations which constitute assertive sentences. The work will be based on recent results concerning the structure of information. These results have shown that each sentence of a language is derived, in a regular and computable way, from a series of more simple statements. It has also been shown that these simple statements contain all of the information which is contained in the original sentence. Thus, the information content of any statement can be derived from the information content of the underlying structures of which the more complex statement can be reduced. The key point is, however, that the derivation of these simple statements do not depend upon human judgment but can be accomplished by a computer program for grammatical decomposition.

The fact that just a few types of successions and combinations of these simple statements suffice for all language-borne information will enable this research to provide insights essential to a theory of information. From the standpoint of applications, the fact that the relationship of information to syntax has been established means that it will be possible for a computer to operate on informational content as opposed to merely word processing. In other words, the development of a systematization of factual information on the basis of a computable formulation of language structure should be of major service toward computer processing of the actual information contained in factual, and especially in scientific, material.

#121

Alternative Representations of Information in Microeconomic Systems. *University of Pennsylvania,* Department of Economics, 409 Franklin Building, 3451 Walnut Street, Philadelphia, PA 19104. Beth Elaine Allen. $77,987 for 24 months beginning July 30, 1979.

To date, little research has been undertaken that defines information in general terms for economic theory. This project uses mathematical economics to analyze some aspects of the role of information in microeconomic theory and examines the relationships between information and microeconomics.

The project is divided into three phases. The definition of information as a forecast will be compared to other ways to model information. The implications of the forecast approach for the notions of similarity of information, more information, and dispersed (or more dispersed) information will be evaluated. The second phase consists of an analysis of a microeconomic model of general equilibrium in which information consists of forecasts and focuses on the sense in which such equilibria can be rational when forecasts are hypothesized to be dispersed. And the third phase will include research on the extension of informational economic equilibrium results to random references in economics and other game-theoretic solution concepts which are relevant to economic theory.

#122

Computer Science and Computer Engineering Research Equipment. *University of Pennsylvania,* Department of Computer and Information Science, Philadelphia, PA 19104. Aravind K. Joshi. $175,100 for 12 months beginning July 2, 1979.

NSF is providing support to the University of Pennsylvania for the purchase of a VAX 11/780 computer system. This system will provide the computing power necessary to conduct several specific research projects in: (1) natural language processing, (2) applicative languages, (3) distributed data bases. The equipment is not intended for general purpose computation, but will be a dedicated facility devoted to specific research projects.

#123

Research in Natural Language Man-Machine Interaction. *University of Pennsylvania,* Department of Computer and Information Science, 3451 Walnut Street, Philadelphia, PA

19104. Bonnie L. Webber. $83,598 for 24 months beginning August 1, 1979.

An important research question in regard to effective information transfer in co-operative natural language man-machine interactions is how a system should respond when it is unable to deal with a user's request. A maximally helpful response would be one which would not frustrate the user or try to outguess him, but rather would guide the user to a relevant answerable question. Such a response would require that a system (1) understand why the user's request is unanswerable, (2) recognize what it would need for an answerable request or what answerable requests are in some useful way related to the given one, and (3) communicate to the user both the problem and its allowable remedies in a way that is both effective and efficient.

This project will investigate these requirements in the context of linguistic notions of repair and computational notions of meta-knowledge. "Repair" is the name given by socialogists and sociolinguists to the set of processes involved in correcting misunderstandings in discourse which may arise as a function of either the speaker or the listener. In this case "meta-knowledge" refers to a system's knowledge of itself as a model of a particular domain and would include the system's knowledge about the extent of its coverage (or lack of coverage) of that domain.

#124

Completions of Relations and Embedded Join Dependencies in Relational Databases: A Study Using Boolean Matrices. *Pennsylvania State University*, Department of Computer Science, 207 Old Main, University Park, PA 16802. Karen H. Chase. $36,630 for 12 months beginning June 26, 1980.

Database management systems are being built around increasingly large and complex databases. To provide flexible and effective responses to the storage, update, and retrieval demands placed upon them, these systems must be designed with a well-organized database model. Such models are provided by the relational approach to database design, a powerful method of representing data at the logical level.

In the relational approach, a key role is played by various semantic constraints which induce so-called data dependencies (e.g., a given plane on a given flight has a unique first pilot). Such dependencies make it desirable to replace complex tables by smaller tables which comprise independent units of information and are especially practical for updates. The study of dependencies is an area of great theoretical and practical interest.

The purpose of the proposed work is to extend the knowledge of these data dependencies in three areas. The first goal is to advance the study of embedded join dependencies by using Boolean matrices. Specifically, the following problem will be addressed: given a set of dependencies, find an algorithm to determine additionally implied dependencies. A second goal is to determine when a relation can be represented by another relation satisfying data dependencies known to exist in the database. Finally, it is proposed to consider the consequences of using these substitute representations and to devise algorithms for applying them to a database.

#125

Training Modules for Use of Scientific and Technical Information Services—Phase II. *University of Pittsburgh*, Graduate School of Library and Information Sciences, Pittsburgh, PA 15260. Elaine Caruso. $67,875 for 12 months beginning February 23, 1978.

Currently, because of the non-standardization existing among computer-stored data bases and the retrieval systems providing access to them, placing users of these systems dirctly on-line, without extensive training, is impractical. Most user training aids now available are system-specific and directed towards information intermediaries, such as librarians, rather than to scientists and engineers or others with technical information needs.

This research, initiated in FY 1976, under SIS76-09538, will continue the design, development, and evaluation of training materials suitable for a wide range of information users. They provide step-by-step on-line instruction aimed at upgrading users' ability to identify data bases suitable for their needs and access them through appropriate search systems. Computer-assisted learning sequences and simulated versions will concentrate on evaluation of the learning and practice modules developed and insurance of their transportability to other computer environments. Results will be reported in the literature and the training packages developed will be made available to the information and education community.

#126

A Survey of Manpower Requirements for Scientific and Technical Communication. *University of Pittsburgh*, Interdisciplinary Department of Information Science, Pittsburgh, PA 15260. Anthony Debons. $101,417 for 7 months beginning June 26, 1978.

The purpose of this research is to determine which communication functions are performed in all sectors of the economy, and the number involved in each function. Two major tasks are described: (1) to develop a taxonomy which will permit the delineation of information work on the basis of functions performed, and will facilitate the classification of occupations in terms of their information function content; and (2) to conduct a survey of the number of persons (and their costs) engaged in information functions at a professional level.

#127

A Survey of Manpower Requirements for Scientific and Technical Communication. *University of Pittsburgh,* Interdisciplinary Department of Information Science, Pittsburgh, PA 15260. Anthony Debons. $20,440 for 12 months beginning April 2, 1979.

This is a supplement to the previous continuing grant described above.

#128

A Survey of Manpower Requirements for Scientific and Technical Communication. *University of Pittsburgh,* Interdisciplinary Department of Information Science, Pittsburgh, PA 15260. Anthony Debons. $126,874 for 12 months beginning June 20, 1979.

This is the final supplement to a continuing grant described above.

#129

Support for U.S. Participation in NATO Advanced Study Institute in Information Science. *University of Pittsburgh,* Interdisciplinary Department of Information Science, Pittsburgh, PA 15260. Anthony Debons. $15,189 for 12 months beginning August 23, 1979.

The Scientific Affairs Division, Science Committee, of the North Atlantic Treaty Organization has approved a grant for the conduct of the Third Advanced Study Institute in Information Science to be held in Crete, Greece, August 1-11, 1978. The Institute will detail the foundations of information science as they relate specifically to the design of information systems. The purpose of this award is to enable eight U.S. scholars, primarily young researchers, to participate. It will also partially support preparation of the Institute's proceedings, as well as other documents to ensure wide dissemination of the results.

#130

The Economics of Information Transfer Using Resource Sharing Networks—Network Functions. *University of Pittsburgh,* Communications Program, Office of the Director, Pittsburgh, PA 15260. Allen Kent. $85,900 for 18 months beginning January 10, 1978.

The objective of this proposal is to obtain a better understanding and assessment of the functions of library resource sharing networks and to identify and define those functions that may be relevant to the eventual development of a model of such resource sharing networks.

This research effort is designed to provide library managers and planners with information that will assist them in determining whether it is to their advantage to participate in a resource sharing network.

#131

Economics of Information Transfer Using Resource Sharing Networks—Network Modeling Simulation. *University of Pittsburgh,* Office of Research, Sponsored Projects Administration, 1028 Cathedral of Learning, Pittsburgh, PA 15260. Allen Kent. $233,496 for 30 months beginning August 13, 1979.

The objectives of this research are to: (1) construct a simulation model capable of accepting data describing network characteristics, functions, rates of activity, financial resources and policies; (2) produce output data that would provide comparative performance measures for alternative choices of resource sharing networks; and (3) prepare and validate a metamodel for resource sharing networks. This model would identify the common parameters of networks, independent of network type, operations, and type of materials.

This research is designed to provide library and information center managers and planners with a tool that will assist them in determining whether it is to their advantage to participate in a resource sharing network.

#132

Research on Error-Correcting Techniques in Automated Information Processing. *Purdue University,* School of Electrical Engineering, West Lafayette, IN 47907. King-Sun Fu. $151,794 for 36 months beginning November 5, 1979.

This project will develop probabilistic analyses to correct errors in information. Error-correcting is important because in practical applications, a certain amount of noise or distortion occurs in the communication, storage, and retrieval of information or in the measurement or processing of physical patterns and languages.

Comprehensive study of such matters requires abstract formulation of information pattern structures and processes. The proposed work utilizes formal languages, a leading example of such an abstraction. They express certain aspects of the syntax (structure) of both human and computer languages and permit the systematic study of processing principles, including parsing, that is, the determination of whether or not a given sequence of words is a possible sentence and, if so, what its structure is.

The proposed approach to recognition of patterns and errors in information exploits an analogy between the structure of patterns and the syntax of language. It thereby seeks to take important results already obtained in formal languages and adapt them to apply in the domain of pattern recognition. The investigator has already had considerable success in this area and proposes to broaden the class of formal languages involved.

Specifically, and in technical terms, this project will extend previous results on stochastic error-correcting parsing for context-free string languages to context-sensitive string languages, languages generated by attributed grammars, and tree languages. Cluster analysis of syntactic information represented by strings and trees also will be studied.

#133

Analysis and Application of Information Productivity Models. *Purdue University,* West Lafayette, IN 47907. Ferdinand F. Leimkuhler. $108,099 for 24 months beginning January 16, 1980.

This study is aimed at clarifying the logical properties of productivity models, developing statistical procedures for analyzing productivity data and demonstrating useful applications of these models to solving problems associated with the design and control of information systems.

The researchers will make a systematic comparison of the various mathematical models that have been proposed to describe the productivity or usefulness of items of information. The basis for these comparisons will be the models' logical or conceptual differences, their usefulness for making statistical inferences about productivity characteristics from sample data, and their usefulness for improving the planning and operation of information systems.

#134

Analysis and Application of Information Productivity Models. *Purdue University,* West Lafayette, IN 47907. Ferdinand F. Leimkuhler. $1,819 for 0 months beginning June 6, 1980.

Supplement to NSF Grant IST79-11893

#135

A Reliability Theoretic Construct for Document Usage and Information Flow. *Rensselaer Polytechnic Institute,* School of Management, Troy, NY 12181. Pasquale Sullo and William A. Wallace. $72,800 for 12 months beginning June 22, 1979.

The research is to develop an analytical model, with associated algorithms, of information flow networks that utilizes the probabilistic framework of reliability theory. Document usage and information flow networks previously developed will be analyzed from a reliability theoretic approach and measures of systems reliability will be determined using existing computer algorithms. The applicability of this approach will be tested by using "typical" networks and performing sensitivity analyses.

The prescriptive nature of the research will permit one to use the results to show how a data procedure can increase the effectiveness of information transfer by identifying the information flow network, determining the probability of information flow between the network nodes, finding measures of system reliability and performing sensitivity analyses to identify the critical components of the system. The result will be a closer congruence of the objectives of the data producer and the requirements of the user.

#136

Information Processing of Logographic and Alphabetic Writing Symbols. *Rensselaer Polytechnic Institute,* Department of Psychology, Troy, NY 12181. Yao-chung Tsao. $27,850 for 12 months beginning August 6, 1979.

As early as 1961 research indicated that the two hemispheres of the human brain function differently. It is now generally accepted that the dominant hemisphere is more efficient in processing verbal material while the non-dominant hemisphere process non-verbal (visuo-spatial) material more efficiently. This project will study information processing with respect to some aspects of hemispheric asymmetry by comparing the processing of two different types of writing symbols—alphabetic (English) and logographic (Chinese). It has been suggested that Chinese characters may be processed initially as pictorial stimuli (by the non-dominant hemisphere) and later, through use and exposure, be processed analytically in the dominant hemisphere.

The general purpose of this project is to investigate the differential information processing of Chinese and English words in relation to brain hemisphere functions. Chinese-English bilinguals will serve as subjects and be tested in each language for three conditions: interference, reading, and color naming. Better understanding of such cognitive processes is expected to contribute to the design of more efficient man-machine interfaces for computer-based information processing systems.

#137

Recognizing Intent in Natural Language. *University of Rochester,* Computer Science Department, 617 Math Science Building, Rochester, NY 14627. James F. Allen. $80,000 for 24 months beginning August 15, 1980.

The proposed research involves the investigation of natural language processing techniques and their use in the design of "intelligent" information retrieval systems. Proper semantic interpretation of [illegible] queries often relies on assumptions about the speaker's goals. This project will examine the ways in which these goals are contextually constrained, inferentially derived, and ultimately applied to understanding processes. A set of plan recognition heuristics will be developed, tested and evaluated.

More specifically, the research aims to develop a formalism for recognizing plans as they occur and can be analyzed in a dialogue. This formalism, a set of plan recognition heuristics, will be implemented in a system for simple exchanges, and then extended and evaluated in a system of long and more complex dialogues.

It is concerned with parsing real inputs to conversation that may not be syntactically well-formed, and devising strategies to efficiently recognize a speaker's plan.

#138

Services in Industry for the Analysis and Evaluation of Information. *Rutgers University*, New Brunswick, NJ 07012. Homer J. Hall. $49,624 for 24 months beginning August 30, 1978.

This project is designed to examine the role of information analysis in the information transfer cycle in industry. Several tasks are identified: (1) define and describe information analysis in industry; (2) provide a critical review of conditions needed for information analysis activities to succeed; (3) make statistical correlations based on the attitude of users; and (4) develop a guide for the potential user and discuss implications for information policy.

#139

Design for Evaluating the Centrality of Basic Information Science Research to Research Users in Organizations. *University of Southern California*, Annenberg School of Communication, Los Angeles, CA 90007. Thomas H. Martin and James A. Danowski. $28,589 for 6 months beginning August 23, 1978.

This six-month award provides for the development of a design to carry out an evaluation of the information science research program. The approach includes: (1) identification of key individuals representative of the information science community; and (2) the development of a set of techniques to assess the importance of various information science research topics of these key individuals. Products of the award will include: (1) specification of the population of organizations containing users of information science research; and (2) preliminary instrument design for isolating key individuals and for assessing the importance of various information science research projects to them.

Several design studies, in addition to this one, are being conducted concurrently. Phase II awards will implement the most promising evaluation designs produced.

#140

Economics of Computer Communication Networks. *Stanford University*, Engineering-Economic Systems Department, Stanford, CA 94305. Donald A. Dunn. $150,000 for 12 months beginning August 3, 1978.

The research program in the economics of computer-communication networks and computer-based information services will be undertaken. This program would continue research carried on during the past five years under a predecessor grant. Studies will be made of the role of information in the U.S. economy, networks as information markets, and various aspects of the production and consumption of information services. These studies will focus on the economics of the information services and products that computer networks make possible.

The major theme of the work under this grant will be to continue to develop policy-oriented models of the information sector of the economy. Projects on the special nature of information as a product, on information service markets, on the distribution and role of information in non-market economics and regulated markets, on the special features of international information markets, and on the structure and behavior of the information sector of the U.S. economy are proposed.

This is continuation of work begun under project MCS78-07973 A03 and MCS78-02272.

#141

Economics of Computer Communication Networks. *Stanford University*, Engineering-Economic Systems Department, Stanford, CA 94305. Donald A. Dunn. $157,172 for 12 months beginning July 1, 1979.

Supplement to the award described above.

#142

Economics of Computer Communication Networks. *Stanford University*, Engineering-Economic Systems Department, Stanford, CA 94305. Donald A. Dunn. $166,458 for 12 months beginning August 1, 1980.

Supplement to award described above.

#143

Theoretical Basis for the Development of Document Preparation Systems. *Stanford University*, Stanford, CA 94305. Donald E. Knuth. $338,707 for 24 months beginning January 23, 1980.

This proposal addresses problems related to the generation, representation, and display of the textual and graphical information that constitutes a "document".

Five main topics are considered for research: (i) Composition of graphical information (creation of figures); (ii) Specification of the layout of a document onto two-dimensional media and optimization of such a process; (iii) Exploration of unifying models of the Document Preparation Process; (iv) Transfer of information between user and the system; and (v) Some computer-related aspects of Document Preparation Systems.

The proposed methodology emphasizes the description of problems, methods, and results using an information model common to all such problems and independent of particular devices or computer systems. The general thrust of the research will be to find a satisfactory scientific foundation for problems that have only been addressed and/or solved for very particular applications. In this way the theoretical methods and results as well as the prototype implementations that are the expected outcome of this proposal will provide a basis for much more powerful, general and easy-to-use Document Preparation Systems.

#144

Developing a Model and Methodology for Evaluating Practitioner Information Systems: The Case of Pacfornet. *Stanford University*, Institute for Communication Research, Stanford, CA 94305. William Paisley. $65,000 for 12 months beginning September 21, 1978.

The proposed research develops an analytical framework and methods for measuring the effectiveness of information systems. The multiple methods for evaluation are developed within the study through an in-depth assessment of a "practitioner" information system, that is, a system set up to communicate research findings to practitioners within such disciplines as forest research, medicine, agriculture, education, etc.

Results of the study, which uses a variety of survey interview and data analysis techniques to assess both individuals and organizational responses to the system, should result in better understanding of the functional relationships between characteristics of information systems and the characteristics of the individuals and organizations which adopt and use them.

In addition, results to be reported in the information and behavioral science literature should begin to overcome the lack of studies in this area and provide a conceptual framework powerful enough to integrate currently existing narrowly focused assessment studies.

#145

Contextual Algorithms for Text Recognition. *State University of New York*, Department of Computer Science, Buffalo, NY, 14260. Sargur N. Srihari. $55,800 for 24 months beginning July 10, 1980.

Text recognition and processing is an important research area of information system design. Conceptually-driven (top-down) methods, data-driven (bottom-up) methods and hybrid methods have been developed to deal with contextual problems in text recognition. The objective of this proposal is to test and develop top-down and bottom-up algorithms for using contextual knowledge in the automatic recognition of patterns of information in text. A process organization that will allow the use of diverse sources of knowledge will also be studied.

The proposed experiments will employ bottom-up contextual information, such as letter- and work-conditional probabilities, and top-down contextual information in the form of dictionaries and knowledge about the text source. The aim is to develop a dynamic system for interpreting text correctly and for automatic correction of errors and identification of patterns of information.

#146

The Relational Theory of Databases. *State University of New York*, Computer Science Department, Stony Brook, NY 11794. David Maier. $89,655 for 24 months beginning August 8, 1979.

The purpose of this research is to extend the theory of the relational data base as well as the model itself. The relational model is one of the easiest of the major database models for the non-technical user to grasp and also enjoys a strong theoretical foundation. In a relational database, all data are contained in tables where each column is labelled with an attribute such as name, age, salary etc. The rows of the table which contain the actual values for each of the attributes are called tuples. Thus, reading across the three columns described above, each row would contain a person's name, his age, and salary. The tables are called relations and a relational database is a set of relations.

There are certain types of semantic information that are not currently incorporated in the model and this research will attempt to structure it in a format which the database designer can apply. An effort will also be made to extend the relational model in order to enable relational databases to become more self-descriptive. This would permit more types of queries to be answered as well as making modifications of the database format easier.

#147

Montague Grammar and Database. *State University of New York*, Department of Computer Science, Long Island, NY 11794. David S. Warren. $79,902 for 24 months beginning August 6, 1980.

It is proposed to relate a linguistic theory and a database theory, each of them the subject of intense current interest in its own field.

The linguistic theory, Montague grammar, uses the principles of mathematical logic to specify syntax of a fragment of English, intensional models, and an interpretation function from phrases in the syntax to objects in the model. A key problem here is to specify a model theory complex enough to represent the real world as viewed by natural language.

The database theory, that of relational databases, also must represent a part of the real world. The representation here should be readily subject to various manipulations and interrogations for different user purposes.

The research plan has two phases, corresponding to two directions of approach. The first begins with a Montague fragment of English and a corresponding intensional model. That model will make it possible to explore how operations defined for relational databases can be used to interpret formulas of the intensional logic. The second phase will deal with the converse problem, starting with a relational database and investigating what Montague grammar and intentional model that database determines. Such issues as the use of the Montague fragment as a query language will then be approachable.

#148

An Evaluation of Factors Affecting Document Ranking by Information Retrieval/Systems. *Syracuse University*, Syracuse, NY 13210. Michael J. McGill. $67,857 for 12 months beginning May 9, 1978.

Information systems attempt to supply lists of documents which appear to be responsive to the user's area of interest. In the case of manual systems it is possible to negotiate your search and narrow down the documents which you will eventually receive by discussions with someone such as a reference librarian. But the introduction of computerized systems has made the process somewhat less personal. One solution has been to rank the output of these systems on the criterion of relevance to the user's query. Ranked outputs attempt to provide the user with information indicating that the closer a document is to the beginning of the output list, the more likely it is to be relevant to his query.

A ranking algorithm defines an ordering on a set of documents, ordering the documents according to their degree of similarity to a query. In the simplest case the retrieved set of documents is ordered into two classes; those satisfying the retrieval criteria and those not satisfying the criteria. More complicated ranking algorithms may provide orderings of greater detail, creating up to N classes of output, where N is the number of documents retrieved.

Many different ranking algorithms have been identified, but little empirical work has been done to analyze the effectiveness and efficiency of those algorithms. This project will examine some three hundred seventy-five ranking algorithms under various environments. The goals are to assess the performance-effectiveness of various terms weighting schemes and similarity measures and to determine the relative costs of implementing and using these algorithms.

#149

A Study of the Impact of Representations in Information Retrieval Systems. *Syracuse University*, School of Information Studies, 113 Euclid Avenue, Syracuse, NY 13210. Michael J. McGill. $155,613 for 24 months beginning February 5, 1980.

Information systems are structured by the way information items are represented; this in turn determines the effectiveness and efficiency of the systems. This project will study different representation schemes to determine the impact of representation on the retrieval characteristics, and suggest performance limits for an information system given a specific representation. At least seven representations will be investigated: (1) title word stems; (2) title-abstract word stems; (3) abstract word stems; (4) controlled vocabulary; (5) classification codes; (6) free indexing terms; and (7) free index stems.

The study will take place over a period of two years and include both a theoretical and an experimental component. The experimental study will emphasize the isolation and testing of specific system components of an information retrieval system. The investigator will isolate, control and vary levels of variables so that main effects can be studied. The theoretical component consists of an iterative process for model building and testing, and formalizing hypotheses to predict behavior of the information systems.

#150

Representation, Retrieval, and Reorganization of Conceptual Information. *University of Texas*, Department of Psychology, Austin, TX 78712. Jane Perlmutter. $63,952 for 24 months beginning July 25, 1979.

This project will explore alternative ways of representing and processing information using techniques developed in cognitive psychology and in computer science, specifically artificial intelligence. It would result in general principles about the structure and processing of information, and provide formal analyses of baseline information processing systems against which the human system can be compared.

A working model of representation, retrieval, and reorganization of conceptual knowledge will be developed. This model should be consistent with psychological constraints and could provide new insights into human memory and cognition.

#151

Individual User Modeling in Computer Systems. *University of Texas*, Austin, TX 78712. Elaine Rich. $56,315 for 24 months beginning August 4, 1980.

The aim of this proposal is to investigate the elements and structure of information retrieval system which contain models of individual users. The proposed work involves a set of stereotypes in a hierarchical arrangement and builds on previous research conducted by the Principal Investigator to develop a system called "Grundy." Stereotypes are similar to other knowledge structures, such as scripts, frames and schemas, and provide a way to represent predictable structures of characteristics and events. Emphasis on these knowledge structures has already resulted in a variety of advances in artificial intelligence. The proposed study will extend principles of knowledge structures and stereotypes to information retrieval systems about food and nutrition.

The project involves developing, storing and mapping knowledge representations of users in a computerized retrieval system. User traits will be modelled as simple linear facets and represented on a linear numerical scale. A set of stereotypes will be developed to depict users of a generalized food and nutrition information system. Conceptual domain in the data base will be designed to match the interests, nutritional requirements, etc. of various stereotypes.

#152

Innovations in Engineering Publications. *Tymshare, Inc.*, 20705 Valley Green Drive, Cupertino, CA 95014. Elizabeth K. Michael. $3,695 for 24 months beginning December 15, 1978.

This project will test an innovative electronic alternative to the traditional paper-based communication of scientific and technical information. It concentrates on information which loses value over time; therefore, rapid dissemination is very important. Engineers in several aerospace organizations are provided with facilities for on-line access to a data base created from inputs of relevant, "time-sensitive" information. Several publishers are participating by converting their regular publication material into the electronic form.

Data on the usefulness of the system will be gathered by pre-test and post-test surveys administered to the subjects and to a control group; and by real time, on-line recording of the engineers' use behavior.

The tests are designed to validate a set of hypotheses regarding the time value of information, reduction of information overload, increase in usefulness of information resources through electronic handling, and improved interaction between the publishers and users.

This is a continuation of work begun under Grant IST77-19308.

#153

Workshop to Compare and Synthesize Evaluations of Computer-Based Information Exchange. *Upsala College*, East Orange, NJ 07019. Starr Roxanne Hiltz. $43,711 for 6 months beginning June 10, 1980.

Investigators of computer-based information exchange will hold a three-day workshop to compare and synthesize findings which they have obtained in twelve separate research projects. Nine of these projects were based on EIES (Electronic Information Exchange System), a research facility developed under NSF sponsorship. The other projects employed PLANET, NLS/AUGMENT, and a system under development at Bell Northern Research, respectively. A fifth system, CONFER, and an electronic mail system, HERMES, are also familiar to some of the participants. The report of this project will summarize each of the separate research projects and include three synthesis chapters. Most of the work of organizing and writing the report will be done by means of EIES. The three-day face-to-face meeting will supplement the participants' computer-based communication and help them to focus their efforts on key aspects of the research under consideration.

#154

Subgraph Identification Using Associative Techniques with Applications to Information Science and Chemical Structure Investigation. *Wayne State University* Department of Electrical and Computer Engineering and Computer Science, 5050 Cass Avenue, Detroit, MI 48202. Dharma P. Agrawal. $61,470 for 24 months beginning July 30, 1979.

Information which is either initially displayed in graphic format or is subsequently encoded graphically (such as organic chemical compounds) can use subgraph identification for purposes of information analysis and retrieval as well as for other pattern recognition purposes. Subgraph identification consists of determining whether or not one graph appears as a subgraph of another. Thus, subgraph identification techniques can be utilized for substructure searches when dealing with the data base of organic chemical compounds. For example, substructure searching would permit the retrieval of only those compounds which contain a given substructure.

The objective of this proposal is to improve the efficiency of subgraph searching techniques. Instead of the currently used algorithms this research will investigate the effectiveness of using distances between nodes in a graph as heuristics to guide the search.

#155

Lexical Ambiguity Resolution in Discourse. *Wayne State University*, Detroit, MI 48202. Michael K. Tanenhaus. $38,977 for 24 months beginning August 25, 1980.

The widespread existence of words with more than one meaning provides an opportunity to investigate the processes by which listeners integrate an ambiguous word and its context in a sentence to determine the appropriate meaning.

The research will examine the interaction between different types of contextual information, lexical retrieval, and comprehension processes, and will make use of a time delay naming paradigm. Unlike previous methodologies used to study lexical ambiguity, this procedure is sensitive to different stages in processing as well as their time course. The studies are designed to test and elaborate a preliminary model of certain aspects of the human word storage and retrieval process which is expected to aid in the development of advanced speech understanding systems.

#156

Accuracy of Data on Information Exchange by Electronic Network. *West Virginia University*, Morgantown, WV 26506. H. Russell Bernard. $39,403 for 12 months beginning August 14, 1978.

Scientists desire to improve the methodology of information science research, particularly the methodology for identifying and studying information networks. They also desire to develop measures of reliability and validity for network data. This project assesses the nature and degree of human error in data gathering by sociometric techniques from information networks. It will also assess the quality of the data that are collected by providing a comparison of sociometric survey data with electronically recorded logs of social interactions. The research will be carried out with the cooperation of volunteers from research groups now using the electronic information exchange system at the New Jersey Institute of Technology.

#157

A Study of the Distribution, Processing, and Use of Scientific and Technical Information in Research Organizations. *University of Wisconsin*, Madison, WI 53706. George P. Huber. $91,214 for 12 months beginning January 23, 1980.

This research is concerned with the distribution, processing, and use of information in organizations. There are three major objectives of the research: (1) develop and test a set of models of information distribution, processing and use in organizations; (2) develop and test transferable procedures for mapping information flows in an organization, especially decision-related information flows; and (3) develop and test transferrable procedures for developing and testing diagnostic, predictive models of an organization's information system.

#158

Distributed Data Base Design. *University of Wisconsin*, School of Business Administration, P. O. Box 413, Milwaukee, WI 53201. William D. Haseman. $243,651 for 18 months beginning August 4, 1980.

Because of recent advances in computer technology, many large central computer facilities are being replaced by clusters of smaller computers. New kinds of software are required for the effective use of networks of databases that are geographically distributed.

This project will develop a methodology and model for designing distributed databases. The modelling will be undertaken as a two-stage process involving first the design of the database (local model) and then the distribution of that database (network model). The design will be accomplished using an integer program which minimizes the sum of access and storage costs. The distribution process uses a location model for determining the number of copies of each file and then their locations. Initially these models will be solved sequentially in an interactive manner. Since this will only produce a local optimum, it is then planned to use the knowledge gained to develop a single model which will combine both steps into one. Solution techniques for this model will be developed and both models will be tested using a medical database.

This project was originally funded as NSF Grant IST79-18472 to Carnegie-Mellon University, and was transferred to the University of Wisconsin, Milwaukee when the investigator accepted a faculty position there.

#159

The Critical Review Process: Impact Upon Scientific and Technological Information. *Worcester Polytechnic Institute*, Department of Management, Worcester, MA 01609. Arthur Gerstenfeld and Richard H. Franke. $147,255 for 24 months beginning August 24, 1979.

The major objective of the project is to establish a conceptual and quantitative model which describes the process and use of critical reviews in the development of science and technology. This is to be accomplished through the analysis of a combination of eight hypotheses which treat the functions of critical reviews and particular characteristics of writing and using them. The research will investigate the critical review process in three areas representing physical science (biochemistry), social science (social science of work and organization), and engineering (electric and electronic engineering). Data are to be obtained in each area: (1) using forward tracking of development through major journals; (2) through reference and citation analyses; and (3) using direct contact with critical review writers, users, and non-users. Analytical methods include graphic representation, t-test comparison, cross-sectional analysis, longitudinal analysis using adjustment procedures of econometrics, and replication and model building employing all of these data and techniques. The examining and contrasting of very different information experiences are expected to yield strong evidence regarding effects of the critical review process upon scientific and technological information.

#160

Text Processing Effects and Memory Recall. *Yale University*, Department of Computer Science, P. O. Box 2158, Yale Station, New Haven, CT 06520. Wendy G. Lehnert, Roger C. Schank, Robert P. Abelson. $291,988 for 24 months beginning August 3, 1979.

This research in the field of natural language processing is concerned with the nature of mechanisms which are capable of performing inferences similar to those which comprise human reasoning. Such mechanisms would permit the development of information retrieval systems with response capabilities much more attuned to user knowledge requirements than can be provided by existing systems.

Using techniques of artificial intelligence, this project will investigate the ways that processing at the time of understanding can affect recall behavior in the tasks of question answering and paraphrase production. This investigation will focus on the design, and implementation of process-oriented memory structures in a computer program that understands narratives.

#161

Theoretical Foundations of Natural Language Processing. *Yale University*, Department of Computer Science, P. O. Box 2158, Yale Station, New Haven, CT 06520. Roger C. Schank. $183,785 for 24 months beginning September 17, 1980.

The objective of this research is to determine whether the fundamental constructs of human knowledge representation are invariant across languages. There are many practical implications which would derive from this finding; one of the most obvious being in the area of knowledge-based automatic text translation. Unlike previous attempts at machine translation, which involved elaborate transformations of syntactic structures between each pair of languages, text understanding systems could be developed based upon an intermediary, language-free, meaning representation. The focus of this work will be transformations from natural language forms to conceptual structures and back to natural language. The use of more than one language will allow a check on the important issue to the extent to which the meaning representations are language-independent.

The work will utilize two different computational understanding systems, one of them already well-developed, the other to be further refined. Use of the two systems will make it possible to investigate two very different aspects of the understanding and generation processes. One of the systems is capable of re-expressing complex descriptions of events, because it can represent and manipulate sophisticated aspects of human interaction. The other system has a much more detailed database of domain-specific knowledge. Both of these aspects must be well-developed in order to achieve an adequate language-independent meaning representation.

Subject Index

(Numbers refer to project summaries)

Academic library management, *26, 93*
Administrative support services, *43*
Algorithms
  — Data dependencies, *124*
  — Document searching, *96, 116*
  — On-line library catalog, *116*
  — Ranking, *148*
  — Searching, *19, 41, 50, 57*
  — Sequential searching, *10*
  — Text recognition, *145*
Allocation process, *19, 113*
Alphabetical writing symbols, *136*
American Chemical Society, *2*
American National Standards Institute
  — Committee Z39, *33*
Analysis
  — Industrial information, *138*
  — Information productivity, *133, 134*
  — Intonation, *83*
  — Languages, *28*
  — Microeconomic systems, *121*
  — Speech, *47, 83, 155*
  — Syntax, *132*
Applications of telecommunication systems, *106*
Applied Communication Research, *3, 4*
Approximate grouping, *60*
Approximate reasoning, *14, 15, 60*
Artificial intelligence, *79, 100, 105, 150, 160*
Artificial systems, *22*
Artificially generated distributions, *44*
Assessment
  — Critical review process, *64, 159*
  — Databases, *36, 72, 110, 111*
  — Document delivery, *49, 76, 84, 87*
  — Document retrieval, *66, 148*
  — Electronic information exchange, *38, 85, 119, 153, 156*
  — Information science research, *46, 71, 94, 139, 150*
  — Information services, *7*
  — Information systems, *3, 27, 66, 67, 82, 99, 132, 138*
  — Information transfer, *17, 70, 135, 144, 159*
  — Libraries, *26, 84, 87, 114*
  — Methodologies, *71, 94, 112, 139*
  — Research, *91, 159*
  — Subject switching, *6*
  — Users' behavior, *4, 17, 125*
  — User training, *125*
  — Videotext, *49*
Associative processes, *32*
Augmented English dictionary, *50*
Automated information processing, *132*
Automated knowledge utilization, *39*
Automated library systems, *115, 116*
Automated natural language generator, *48*
Automated subject switching, *6, 5*
Automated vocabulary switching, *5*
Automatic classification, *107*
Automatic detection of overlap, *56*
Automatic detection of spelling errors, *2*
Automatic identification of knowledge, *22*
Automatic indexing, *58*
Automatic information retrieval, *31*
Automatic pattern recognition, *145*
Automatic phonetic analysis, *47*
Automatic speech understanding, *83*
Automatic text translation, *161*

Basic language structures, *28*
Battelle Memorial Institute, *6, 5*
Bayesian retrieval system, *13*
Bayesian statistics, *13*
Behavioral variables, *79*
Bell Northern Research, *153*
Bethune-Cookman College, *7*
Bibliographic citations, *65*
Bibliographic search and retrieval, *37*
Bibliometrics, *62, 65*
Binary index terms, *81*
Boolean matrices, *124*
Boston University, *8*
Brandeis University, *9*
Brooklyn College, *10*
Brown University, *11*
Bureau of Social Science Research, *12*

Carnegie-Mellon University, *22*
Case Western Reserve University, *24*
Chemical Abstracts Service, *2*
Chemical information, *7, 154*
Citations, *62, 65, 73, 75, 159*
Citing statements, *75*
City University of New York
  — Queens College, *25*
Clarkson College of Technology, *26*
Classification, *2, 24, 107*
Clusters, *62, 65, 91, 132*
Cognitive aids programs, *27*
Cognitive processes, *8, 11, 16, 66, 77, 100*
Cognitive psychology
  — Techniques, *79*
Columbia University, *28, 29*
Communication among scientists, *92*
Communication processes, *24, 25, 105*
Communication systems, *51, 92, 101, 153*
Computational gains, *18*
Computational linguistics, *123*
Computer applications to chemistry, *110, 111*
Computer architecture, *23, 158*
Computer-based information resources, *67, 140*
Computer-based investigations, *109*
Computer communication networks, *140, 142, 153*
Computer conferencing, *38, 61, 85, 98, 101, 102, 119, 153, 156*
Computer languages, *52*
Computer research equipment, *122*
Computer speech processing, *47*
Computerized journals, *75*
Computerized library circulation, *114*
Conceptual information, *150*
Concurrent manipulation of structured data, *118*
Contextual algorithms, *145*
Controlled vocabularies, *6*
Copyright, *69, 70*

Cornell University, *30, 31, 32*
Correction heuristics, *2*
Costs
  — Computerized information, *7, 18, 67, 72*
  — Information retrieval, *32, 53*
  — Libraries, *26, 59*
  — Library acquisitions, *59*
  — STI, *3, 9, 70*
Council of National Library Associations, *33*
Critical review processes, *159*
Critically evaluated data, *97*
Cross-impact analysis of STI, *42*

Data dependencies, *124*
Data retrieval, *30*
Data semantics, *21*
Data structure, *53*
Data Use and Access Laboratories, Inc., *34*
Databases
  — Analysis, *1, 56, 104, 118, 147*
  — Bibliographic, *36, 55, 56, 115*
  — Chemistry, *2, 154*
  — Design, *1, 21, 23, 35, 118, 146, 147, 158*
  — Distributed, *19, 23, 158*
  — Language, *25, 35*
  — Management, *97, 108, 124*
  — Numerical, *97*
  — Social Sciences, *12, 34, 62*
  — Theories, *1, 11, 80, 124, 146*
Decision analysis, *14, 117*
Decision-making processes, *8, 20, 27, 60, 105, 117, 157*
Degradation of accuracy, *40*
Demand for journals, *78*
Devices for the disabled, *61*
Dictionary for information retrieval, *50*
Distributed databases, *19, 23, 158*
Documents
  — Delivery, *84, 86*
  — Preparation, *143*
  — Retrieval, *30, 81, 86*
  — Use, *135*
Drexel University, *36*
Duplicate journal articles, *56*
Dynamic multivariate display, *12*
Dynamics of information, *24*

Economics of information, *3, 9, 70, 113, 117, 128, 130, 131, 142*
Ediometer, *45*
Educom, *37*
Effectiveness of information transfer, *82*
Effects of STI changes on engineers, *44*
Electronic alternative, *70, 152*
Electronic document delivery, *76, 84, 87*
Electronic Industries Association, *38*
Electronic information exchange
  — Operational trials, *38, 40, 61, 85, 98, 119, 153, 156*
  — Test facility, *101, 102*
Electronic interlibrary loans, *84, 87*
Electronic networks, *84, 156*
Electronic text distribution, *49*
Energy information, *6, 5*
Engineering education, *67*
Engineering information, *44, 82, 99, 152*
Equipment for information science research, *22, 122*
Error correcting techniques, *132*
Evaluation proceedures design, *71, 94, 139*

Fact retrieval, *107*
False retrieval, *75*
Federal Library Committee, *76*
Florida State University, *41*
Flow of STI, *95, 99, 104, 135*
Forecasting International, Ltd., *42*
Formal grammar research, *90*
Formal linguistics, *90*
Foundation of information science, *127*
Foundation of measuring theories, *16*
Framework for information science research, *112*
Franklin Institute Research Laboratories, *36, 43*
Full text
  — Retrieval, *75*
  — Searching, *73*
Future changes in STI process, *42, 44*
Fuzzy sets, *60*

Georgia Institute of Technology, *45, 46*
Graph theory mathematics, *24*
Graphics information, *143*
Grundy, *151*

Haskins Laboratories, Inc., *47*
Hermes, *153*
Herner and Co., *76*
Hofstra University, *48*
Human behavior, *54*
Human brain functions, *136, 150*
Human information processes, *25, 77, 103, 105, 161*
Human interaction representation, *161*
Human knowledge representation, *161*
Human visual systems, *88*

Ill formed input, *35*
Illinois Institute of Technology, *58*
  — Research Institute, *49, 50*
Implications of paperless society, *51*
Indexing
  — Binary terms, *81*
  — Theories, *31*
Indiana University, *59, 60*
Indicators
  — Articles, *109*
  — Information science, *68, 71, 94, 139*
  — Journals, *56, 59, 69, 78*
  — STI, *63, 69, 70*
  — Scientific and technical literature, *12, 63, 68*
Individual user modelling, *151*
Industrial information, *138*
Inference methods, *14, 15*
Informal information networks, 99
Information
  — Analysis, *117, 133, 134, 137, 154*
  — Assessment, *18, 121, 138*
  — Behavior, *4*
  — Computer-based, *67, 109*
  — Decision making, *8, 27, 60, 117, 157*
  — Dynamics, *24*
  — Economics, *9, 53, 69, 113, 117, 121, 130, 140, 142, 152*
  — Flow, *117*
  — Graphics, *143*
  — Language, *28, 48, 120*
  — Modelling, *25, 54, 81, 96, 103, 121, 142, 143*
  — Processes, *22, 25, 32, 45, 77, 81, 88, 92, 100, 103, 107, 118, 136, 150*
  — Representation, *15, 48, 77, 88, 121, 149, 151*
  — Retrieval, *13, 30, 31, 32, 41, 50, 53, 57, 60, 66, 75, 81, 86, 96, 105, 107, 123, 132, 137, 148, 151, 154*
  — Sharing, *89, 130*
  — Structure, *16, 20, 28, 29, 45, 108, 109, 118, 132*
  — Theories, *20, 22, 46, 117, 120, 121, 128, 150*
  — Use, *44, 54, 117, 135, 152, 157*
Information productivity models, *133, 134*

Information resources for engineering education, *67*
Information role in the economy, *140, 142*
Information science
— Equipment, *122*
— Foundation, *46, 127, 129*
— Indicators, *68, 71, 94, 139*
— Literature, *62*
— Methodology, *105, 156*
— Program, *94, 105*
— Research framework, *105, 112, 139*
— Research users, *40, 139, 152*
Information services
— Computer-based, *7, 67, 72, 89, 115, 140, 142*
— Economics, *142*
— Handicapped users, *61*
— Planning, *98*
— Standards, *33, 34*
— Training, *125*
Information systems
— Analysis, *54, 70*
— Behavioral variables, *79*
— Design, *13, 14, 27, 30, 35, 49, 54, 79, 100, 107, 110, 111, 127, 133, 134, 145*
— Documents, *148*
— Flow, *99*
— Modelling, *1, 17, 18, 20, 66, 133, 134, 149*
— Natural languages, *35, 108*
— Retrieval, *13, 66, 75, 160*
— Users, *40, 66, 123, 135, 160*
Information theory impact, *46*
Information transfer
— Documents, *143*
— Economics, *70, 130, 131*
— Effectiveness, *25, 82, 123, 135*
— Libraries, *98*
— Networks, *6, 38, 85, 101, 119, 130, 156*
Innovations
— Engineering publications, *152*
Innovative Systems Research, Inc., *61*
Institute for Scientific Information, *62, 63, 64, 65*
Intelligent systems, *74, 137*
Intensional logic, *147*
Interaction among publishers, *65*
International Council of Scientific Unions
— Abstracting Board, *95*
International Federation for Documentation, *95*
International information programs, *95*
Intonation, *83*

Johns Hopkins University, *66*
Joint Electronic Device Engineering Council, *38*
Journals
— Citations, *75*
— Computerized text, *73*
— Critical reviews, *159*
— Indicators, *56, 59, 69, 78*

K- d trees, *41*
King Research, Inc., *68, 69, 70, 71, 72*
Knowledge
— Management, *74*
— Organization, *39*
— Representation, *8, 39, 151, 161*
— Structure, *151*
— Transfer, *25*

LUST theorem proving, *80*
Language
— Structure, *15, 28, 29, 120*
— Zellig Harris theory, *29*
Language-borne information, *120*
Learning procedures of artifical systems, *22*
Legislative Resources Network, *119*
Lehigh University, *73, 74, 75*
Lexical ambiguity, *155*
Lexical-semantic relationship, *58*
Libraries, *7, 26, 33, 59, 72, 76, 84, 87, 93, 98, 104, 114, 115, 116, 130*
Library acquisitions, *59*
Library automation, *69, 72, 114, 115*
Library policy, *59*
Library science *105*
Logic-based formal systems, *90*
Logical database design, *21*
Logographic symbols, *136*

Machine-readable files, *34, 55*
Man-machine interface, *54, 55, 66, 123, 136*
Management of knowledge, *74*
Manpower for STI, *126, 128*
Mapping of information, *109*
Market structure in STI, *9*
Markov processes, *54*
Massachusetts Institute of Technology, *82, 83, 84, 85, 86, 87, 88*
Mathematical logic application, *80*
Matrix memory models, *11*
Measurements
— Artificial words, *45*
— Citations, *65*
— Information systems, *60, 144*
— Models, *103*
— Research results, *71*
— Social data, *12*
— Theories, *16*
Memory processes, *11*
Memory recall, *160*
Mental workload research, *85*
Meta knowledge, *123*
Microcomputer standards, *38*
Microeconomic systems, *121*
Minnesota Higher Education Coordinating Board, *93*
Mitre Corporation, *94*
Modelling
— Communication, *24, 70, 92, 144*
— Conceptual knowledge, *150*
— Critical review process, *159*
— Databases, *8, 21, 23, 158*
— Document preparation, *143*
— Formal linguistics, *90*
— Human behavior, *54*
— Information processes, *11, 25, 54, 103*
— Information productivity, *133, 134*
— Information science research, *112*
— Information systems, *1, 8, 17, 18, 55, 72, 131, 142, 143, 144, 149*
— Journals, *78*
— Microeconomic systems, *121*
— Relational theories, *124, 146, 147*
— STI, *42, 69, 135, 157*
— Searching, *41*
— Storage and retrieval, *31, 36, 66, 96*
— Structured information, *77, 151*
— Videotext, *49*
Models of attention
— Evaluation, *100*
Montague grammar, *90, 147*
Multiple attribute retrieval, *41*
Multiple database searching, *56*

N-gram analysis, *2*
NLS-augment, *153*
NTL Institute for Applied Behavioral Science, *112*

Name matching, *115*
Narrative analysis, *160*
National Academy of Sciences/National Research Council, *95, 97*
NATO Advanced Study Institute on Informaion Science, *127, 129*
Natural computation, *88*
Natural language
  — Databases, *35, 90, 108, 147*
  — Processing, *15, 35, 45, 48, 52, 58, 107, 123, 137, 160, 161*
Network resource utilization, *37, 55, 131*
Networks, *8, 19, 23, 37, 55, 76, 84, 87, 89, 91, 99, 116, 130, 131, 135, 140, 142, 156, 158*
New Jersey Institute of Technology, *38, 40, 101, 102, 153*
New York University, *103, 104, 105, 106, 107, 108, 109*
Non-narrative paragraphs, *48*
Normal sentence forms, *29*
Northwestern University, *113*
Numerical Data Advisory Board, *97*

Oberlin College, *114*
Ohio College Library Center, Inc., *115, 116*
Ohio State University, *117*
On-line library catalogs, *116*
On-line knowledge transfer, *39*
Optimal file allocation, *19*
Organization of information, *20*
Organization of knowledge, *39*
Organization of scientific specialties, *92*

PRIM-9 systems, *9*
Paperless communication systems, *51*
Paperless society, *51*
Paragraph patterns, *48*
Parsing *132, 137*
Participation Systems, Inc., *119*
Pattern recognition, *132, 145, 154*
Pattern of STI by engineers, *44, 152*
Peer reviewing, *64, 91*
Pennsylvania State University, *124*
Perfect match retrieval, *32*
Performance allocation process, *113*
Performance measurement of library operations, *26, 114*
Planet, *153*
Private investment for STI, *3*
Probabilistic analysis, *132*
Probabilistic estimation, *13*
Probabilistic matching, *31, 115*
Probabilistic models, *103*
Procedural learning, *22*
Process-oriented memory structures, *160*
Processing of writing symbols, *136*
Publishing, *65, 104, 152*
Purdue University, *132, 133, 134*

Quantitative data, *97*
Query formulation, *19, 66*

Ranking algorithms, *148*
Recall and relevance, *7, 13, 36, 100, 148, 156, 160*
Recognizing intent in natural languages, *137*
Reference retrieval, *73*
Regional computer networks, *37*
Rehabilitation technology products, *61*
Relational database theory, *1, 124, 146*
Reliability theory; *135*
Rensselaer Polytechnic Institute, *79, 136*
Repair sequences, *123*
Representation of information, *15, 48, 77, 88, 121, 149, 151*
Representation of knowledge, *8, 39, 151, 161*
Research libraries, *51*
Research organizations
  — STI, *157*
Resource allocation in libraries, *26*
Resource sharing, *37, 55, 93, 130, 131*
Reviews, *64*
Robotic systems, *14*
Rutgers University, *138*

STI see Scientific and technical information
Science policy, *63*
Scientific and technical information
  — Administrative support services, *43*
  — Analysis, *9, 43, 70, 92*
  — Changes, *42, 44*
  — Computer-based systems, *7, 67, 72, 89, 115, 140, 142*
  — Critical review processes, *159*
  — Economics, *3, 9, 70, 113, 117, 128, 130, 131, 142*
  — Flow, *95, 99, 104, 135*
  — Indicators, *63, 69, 70*
  — International, *95*
  — Manpower, *128*
  — Surveys, *69*
  — Training modules, *125*
  — Use, *44, 125, 127*
Scientific and technical literature
  — Articles, *109*
  — Indicators, *63, 68*
Scientific communications, *68, 92, 104, 126, 128*
Scientific reviews, *64*
Scientific specialities, *92*
Searching, *5, 7, 10, 41, 57, 73, 86, 96, 110, 111, 148, 154*
Sequential search algorithms, *10*
Serial collections, *93*
Sharing technical information, *89*
Sign shapes, *45*
Simultaneous equation models, *78*
Social networks, *40*
Social Sciences Citation Index, *62*
Social sciences research
  — Databases, *12, 34, 62*
Social system of science, *63*
Sociolinguistics, *123*
Sociometric surveys, *40*
Spark-Jones and Robertson, *81*
Spatially structured information, *77*
Speech
  — Analysis, *47, 83, 155*
  — Intonation, *83*
Speech understanding systems, *83*
Spelling errors detection, *2*
Sperry Univac, *35*
Standards
  — Information services, *33, 34*
  — Libraries, *33*
  — Microcomputers, *38*
  — Telecommunications, *38*
Stanford Research Institute, *152*
Stanford University, *140, 142, 143, 144*
State University of New York
  — Buffalo, *145*
  — Stony Brook, *146, 147*
Stochastic error-correcting, *132*
Storage and retrieval
  — Accessing, *53, 57*
  — Error-correcting *132*
  — Modelling, *31, 36, 66, 96*
  — Natural languages, *52*
  — Words, *155*
Structured data, *53, 118*
Subject switching, *6, 5*
Syllable hypothesis, *47*

Symbolic transportation parameters, *47*
Syntactic structure of information, *45*, *132*
Syntax analysis, *132*
Synthetic language, *15*
Syracuse University, *148*, *149*
Systems theory, *105*

Technology impact on research libraries, *51*
Telecommunications, *106*
Term matching, *31*, *32*
Test facility, *38*, *85*, *101*, *119*, *156*
Text processing, *145*, *160*
Text recognition, *145*
The Futures Group, *44*
Theorem proving application, *80*
Theoretical issues in natural languages processing, *52*
Theory of information flow and analysis, *117*
Training
  — Materials for STI users, *125*
  — Modelling human behavior, *54*
Tree-structured information, *41*
Trends in STI, *42*, *44*
Tymshare, Inc., *152*

Understanding speech, *83*
United States
  — Air Force Department, *1*
  — Commerce Department, *96*, *97*
  — Library of Congress, *76*
  — National Commission on Libraries and Information Services, *98*
  — Navy Department, *99*
University of California
  — Berkeley, *13*, *14*, *15*
  — Irvine, *16*
  — Los Angeles, *17*, *18*, *19*, *20*, *21*
University of Colorado, *27*
University of Delaware, *35*
University of Florida, *39*, *40*
University of Illinois, *51*, *52*, *53*, *54*, *55*, *56*, *57*
University of Kansas, *67*
University of Louisville, *77*
University of Lowell, *78*
University of Maryland, *79*, *80*
University of Massachusetts, *81*
University of Michigan, *89*, *90*, *91*, *92*
University of Nebraska, *100*
University of North Carolina, *110*, *111*
University of Oregon, *118*
University of Pennsylvania, *29*, *120*, *121*, *122*, *123*
University of Pittsburgh, *125*, *126*, *127*, *128*, *129*, *130*, *131*
University of Rochester, *137*
University of Southern California, *139*
University of Texas
  — Austin, *150*, *151*
University of Wisconsin
  — Madison, *157*
  — Milwaukee, *158*
Upsala College, *153*
Usefulness of citation indexes, *73*
User requirements, *64*

Van Rysbergen, *81*
Vector processing, *31*
Videotext, *49*
Visual information, *100*
Vocabulary
  — Design, *58*, *5*
  — Information retrieval, *50*
  — Lexical-semantic relations, *2*, *58*
  — Switching, *6*, *5*

Wayne State University, *154*, *155*
West Virginia University, *40*, *156*
White House Conference
  — Library and Information Services, *98*
Worcester Polytechnic Institute, *159*
Word structure, *2*, *28*
Writing symbols, *136*

Yale University, *160*, *161*

Zellig Harris theory, *29*

END

8-23-84